PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, C12Q 1/68, A61K 38/17, C07K 16/18

A2

(11) International Publication Number: **WO 00/28032**

(43) International Publication Date: 18 May 2000 (18.05.00)

(21) International Application Number: PCT/US99/26742

(22) International Filing Date: 12 November 1999 (12.11.99)

(30) Priority Data:

09/191,280	12 November 1998 (12.11.98)	US
Not furnished	12 November 1998 (12.11.98)	US
09/206,647	7 December 1998 (07.12.98)	US
Not furnished	7 December 1998 (07.12.98)	US
60/123,404	8 March 1999 (08.03.99)	US

(63) Related by Continuation (CON) or Continuation-In-Part (CIP) to Earlier Applications

US	09/191,280 (CIP)
Filed on	12 November 1998 (12.11.98)
US	Not furnished (CIP)
Filed on	12 November 1998 (12.11.98)
US	09/206,647 (CIP)
Filed on	7 December 1998 (07.12.98)
US	Not furnished (CIP)
Filed on	7 December 1998 (07.12.98)
US	60/123,404 (CIP)
Filed on	8 March 1999 (08.03.99)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Moutain View, CA 94043 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). AU-YOUNG, Janice [US/US]; 1419 Kains Avenue, Berkeley, CA 94702 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN CELL SURFACE RECEPTOR PROTEINS

(57) Abstract

The invention provides human cell surface receptor proteins (HCSRP) and polynucleotides which identify and encode HCSRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of HCSRP.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

HUMAN CELL SURFACE RECEPTOR PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human cell surface receptor proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative disorders, immune system disorders, infections, and neuronal disorders.

BACKGROUND OF THE INVENTION

The term receptor describes proteins that specifically recognize other molecules. The category is broad and includes proteins with a variety of functions. The bulk of receptors are cell surface proteins which bind extracellular ligands and produce cellular responses in the areas of growth, differentiation, endocytosis, exocytosis and immune response. Central to the function of cell surface receptors is the capacity to adhere or bind to other proteins or ligands through special functional domains.

Cell surface receptors are typically integral plasma membrane proteins. These receptors recognize hormones such as catecholamines; peptide hormones; growth and differentiation factors; small peptide factors such as thyrotropin releasing hormone; galanin, somatostatin, and tachykinins; and circulatory system-borne signaling molecules. Cell surface receptors on immune system cells recognize antigens, antibodies, and major histocompatibility complex (MHC)-bound peptides. Other cell surface receptors bind ligands to be internalized by the cell. This receptor-mediated endocytosis functions in the uptake of low density lipoproteins (LDL), transferrin, glucose- or mannose-terminal glycoproteins, galactose-terminal glycoproteins, immunoglobulins, phosphovitellogenins, fibrin, proteinase-inhibitor complexes, plasminogen activators, and thrombospondin (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, p. 723; and Mikhailenko, I. et al. (1997) J. Biol. Chem. 272:6784-6791).

Many growth factor receptors, including receptors for epidermal growth factor, platelet-derived growth factor, fibroblast growth factor, as well as the growth modulator α-thrombin, contain intrinsic protein kinase activities. When growth factor binds to the receptor, it triggers the autophosphorylation of a serine, threonine, or tyrosine residue on the receptor. These phosphorylated sites are recognition sites for the binding of other cytoplasmic signaling proteins. These proteins participate in signaling pathways that eventually link the initial receptor activation at the cell surface to the activation of a specific intracellular target molecule. In the case of tyrosine residue autophosphorylation, these signaling proteins contain a common domain referred to as a src homology 2 (SH2) domain. SH2 domains are found in a variety of signaling molecules and oncogenic proteins

such as phospholipase C-γ, Ras GTPase activating protein, and pp60^{c-src} (Lowenstein, E.J. et al. (1992) Cell 70:431-442).

G-protein coupled receptors (GPCRs)

G-protein coupled receptors (GPCRs) are integral membrane proteins characterized by the presence of seven hydrophobic transmembrane domains which span the plasma membrane and form a bundle of antiparallel alpha (α) helices. These proteins range in size from under 400 to over 1000 amino acids (Strosberg, A.D. (1991) Eur. J. Biochem. 196:1-10; Coughlin, S.R. (1994) Curr. Opin. Cell Biol. 6:191-197). The amino-terminus of the GPCR is extracellular, of variable length and often glycosylated; the carboxy-terminus is cytoplasmic and generally phosphorylated. Extracellular loops of the GPCR alternate with intracellular loops and link the transmembrane domains. The most conserved domains of GPCRs are the transmembrane domains and the first two cytoplasmic loops. The transmembrane domains account for structural and functional features of the receptor. In most cases, the bundle of alpha helices forms a binding pocket. In addition, the extracellular N-terminal segment or one or more of the three extracellular loops may also participate in ligand binding. Ligand binding activates the receptor by inducing a conformational change in intracellular portions of the receptor. The activated receptor, in turn, interacts with an intracellular heterotrimeric guanine nucleotide binding (G) protein complex which mediates further intracellular signaling activities, generally the production of second messengers such as cyclic AMP (cAMP), phospholipase C, inositol triphosphate or ion channel proteins (Baldwin, J.M. (1994) Curr. Opin. Cell Biol. 6:180-190).

One group of GPCRs are the rhodopsin-like GPCRs that transmit extracellular signals of diverse natures including hormones, neurotransmitters and light. Rhodopsin is a photosensitive GPCR in the vertebrate eye. Rhodopsin, which defines a conserved subfamily of GPCRs found in animal retinas, is about 350 amino acids in length. In vertebrates, rhodopsin molecules are embedded in membranous stacks found in photoreceptor (rod) cells. Each rhodopsin molecule responds to a photon of light by triggering a decrease in cGMP levels which leads to the closure of plasma membrane sodium channels. In this manner, a visual signal is converted to a neural impulse. Other rhodopsin-like GPCRs are directly involved in responding to neurotransmitters. These GPCRs include the receptors for adrenaline (adrenergic receptors), acetylcholine (muscarinic receptors), adenosine, galanin, and glutamate (N-methyl-D-aspartate/NMDA receptors). (Reviewed in Watson, S. and Arkinstall, S. (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego, CA, pp. 7-9, 19-22, 32-35, 130-131, 214-216, 221-222; Habert-Ortoli, E. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9780-9783.)

The somatostatin receptor type 4 is another example of a rhodopsin-like GPCR. It is one of several high affinity receptors for somatostatin, a tetrapeptide that inhibits the secretion of growth

hormone from the anterior pituitary. Expression of particular somatostatin receptors has been linked to the efficacy of drug therapy in specific endocrine tumors (Kubota. A. et al. (1994) J. Clin. Invest. 93:1321-1325).

Another rhodopsin-like GPCR is the prostanoid EP1 receptor that recognizes prostanoids such as prostaglandin to mediate a variety of physiological functions including cardiovascular and immune responses. EP1 receptors have a role in the contraction and relaxation of smooth muscle and can activate the phosphoinositide pathway (Watson, supra, pp. 239-251). The prostanoid DP receptor is another rhodopsin-like GPCR that is specific for prostaglandin D2 (PGD2). Expression of the DP receptor has been localized to the mammalian brain and eye tissues and upon activation facilitates elevation of intracellular cAMP and Ca2+ mobilization but does not generate inositol 1,4,5-triphosphate (Boie. Y. et al. (1995) J. Biol. Chem. 270:18910-18916; Gerashchenko. D. et al. (1998) J. Neurochem. 71:937-945).

Still another rhodopsin-like GPCR is the endothelin receptor that plays a role in cardiovascular system regulation through endothelins. Endothelins are potent vasoconstrictors that can stimulate cardiac and smooth muscle contraction as well as stimulate secretion in tissues such as kidney, liver and adrenals. Endothelin receptors may have a role in the brain, where they are also found, and there is evidence that endothelins may be associated with pathophysiological conditions such as stress (Watson, supra, pp. 111-116).

The secretin receptor is an example of a unique GPCR that responds to secretin, a peptide hormone that stimulates the secretion of enzymes and ions in the pancreas and small intestine (Watson, supra, pp. 278-283). Secretin receptors are about 450 amino acids in length and are found in the plasma membrane of gastrointestinal cells. Binding of secretin to its receptor stimulates the production of cAMP. An unusual member of the secretin receptor family has been identified from a neuroectodermal cDNA library (Baud. V. et al. (1995) Genomics 26:334-344). This receptor, EMR1 (EGF-like, mucin-like hormone receptor), is 886 amino acids in length and contains six epidermal growth factor (EGF)-like modules at the N-terminus followed by a serine/threonine rich domain. The latter feature is characteristic of mucin-like integral membrane adhesive proteins.

Other GPCRs have been identified which play a role in the immune response. For example, a new subfamily of GPCRs has been identified from a human monocyte (HM) cDNA library (Nomura, H. et al. (1993) Int. Immunol. 5:1239-1249). Most of these GPCRs likely bind to cytokines and other leukocytic signaling molecules. One of these GPCRs, HM74, is particularly unusual in that its N-terminus does not contain N-glycosylation sites.

The thrombin receptors (TRs) have GPCR activity and are activated by the ligand α-thrombin. Through TR-mediated signal transduction pathways, α-thrombin induces production of IL-8 and IL-6

in cultured monocytes and endothelial cells (Johnson, K. et al. (1998) J. Immunol. 160:5130-5135). Conversely, α-thrombin inhibits the action of IL-6, leukemia inhibitory factor, and ciliary neurotrophic factor in chinese hamster lung fibroblasts (Bhat, G.J. et al. (1998) Arch. Biochem. Biophys. 350:307-314). In addition, when α-thrombin binds to the TR it proteolytically cleaves 40 amino acids from the N-terminus of the receptor. The cleaved peptide is termed the thrombin receptor agonist peptide and acts as a tethered ligand for the TR to increase the potency of the thrombin-derived signal (Hou, L. Et al. (1998) J. Periodontal Res. 33:205-211; Johnson, et al. supra).

GPCR mutations, which may cause loss of function or constitutive activation, have been associated with numerous human diseases (Coughlin, supra). For instance, retinitis pigmentosa may arise from mutations in the rhodopsin gene. Parma, J. et al. (1993, Nature 365:649-651) report that somatic activating mutations in the thyrotropin receptor cause hyperfunctioning thyroid adenomas and suggest that certain GPCRs susceptible to constitutive activation may behave as protooncogenes.

Cytokine Receptors

Cytokines comprise a family of signaling molecules that modulate the immune system and the inflammatory response. Cytokines are usually secreted by leukocytes, or white blood cells, in response to injury or infection. However, other tissues are capable of secreting cytokines in response to disease or other physiologic perturbations. Cytokines function as growth and differentiation factors that act primarily on cells of the immune system such as lymphocytes, monocytes, and granulocytes. Like other signaling molecules, cytokines bind to specific plasma membrane receptors and trigger intracellular signal transduction pathways which regulate gene expression, cell proliferation, and cell differentiation.

Erythropoietin (EPO) is an unusual cytokine that is produced not by leukocytes, but instead by the kidney or liver. EPO stimulates erythroid precursors to differentiate into red blood cells. EPO also stimulates the production of platelets. The EPO receptor is a single-pass transmembrane protein of about 500 amino acids, the intracellular domain of which associates with JAK2 kinase. Activated EPO receptor stimulates the phosphorylation activity of JAK2 which triggers gene transcription and mitogenesis. (Reviewed in Callard, R. and Gearing, A. (1994) The Cytokine Facts Book, Academic Press, San Diego, CA, pp.114-118.)

Immunoglobulin Domain Receptors

Immune system and related cell surface receptor proteins have hallmark functional domains (for example the Immunoglobulin (Ig) domain) with similar functionality in a wide array of receptor types. The human immune system is responsible for combating infectious agents contracted from the environment. One critical component of the immune system are highly specialized molecules called immunoglobulins (Ig) or antibodies that can recognize and bind to foreign antigens, thereby

facilitating their elimination. Characteristic features of immunoglobulins include their structural motifs that include regions for membrane attachment, antigen recognition (variable (V) regions), and polymerization. Polymerized immunoglobulins such as glandularly secreted IgA and IgM must undergo transcellular transport, a process mediated by the poly-immunoglobulin (poly-Ig) receptor. The poly-Ig receptor is itself a member of the immunoglobulin superfamily having homology to the variable (V) region of immunoglobulins (Hood, L. et al. (1985) Cell 40:225-229). Like all immunoglobulin superfamily members, poly-Ig receptor is involved in adhesion or binding to other proteins through the conserved immunoglobulin-like domain. This Ig domain is comprised of antiparallel β sheets joined by a disulfide bond in an arrangement called the Ig fold. Members of the Ig superfamily include T-cell receptors, MHC proteins, CD4, CD8, and CD 28 cell surface proteins, and antibodies.

Immunoglobulins, or antibodies, are the central components of the humoral immune response. IgG, the most common class of immunoglobulin in the circulation, can be described in terms of two main functional domains. Antigen recognition is mediated by the Fab (antigen binding fragment) region of the IgG, while effector functions are mediated by the Fc (crystallizable fragment) region. Binding of IgG to an antigen, such as a bacterium, triggers the destruction of the antigen by phagocytic white blood cells, such as macrophages and neutrophils. These cells express cell surface receptors that specifically bind to the IgG Fc region and allow the phagocytic cells to engulf, ingest, and degrade the IgG-bound antigen. The IgG Fc receptors expressed by phagocytic cells are single-pass transmembrane glycoproteins of about 400 amino acids (Sears, D. W. et al. (1990) J. Immunol. 144:371-378). The extracellular portion of the IgG Fc receptor typically contains two or three Ig domains.

T cells play a dual role in the immune system as effectors and regulators, coupling antigen recognition with the transmission of signals that induce cell death in infected cells and stimulate other immune cells. Although T cells collectively recognize a wide range of different antigens, a clonal line of T cells can only recognize a single antigen. Moreover, the antigen must be presented to the T cell receptor (TCR) as a peptide complexed with a major histocompatibility molecule (MHC) on the surface of an antigen-presenting cell. The TCR on most T cells consists of two polypeptide subunits, α and β, which are immunoglobulin-like integral membrane glycoproteins of similar molecular weight. The TCRα and TCRβ subunits have an extracellular domain containing both variable and constant regions, a transmembrane domain that traverses the membrane once, and a short intracellular domain (Saito, H. et al. (1984) Nature 309:757-762). The genes for the TCR subunits are constructed through somatic rearrangement of different gene segments. Interaction of antigen in the proper MHC context with the TCR initiates signaling cascades that induce the proliferation, maturation, and

function of cellular components of the immune system (Weiss. A. (1991) Annu. Rev. Genet. 25: 487-510). Rearrangements in TCR genes and alterations in TCR expression have been noted in lymphomas. leukemias. autoimmune disorders. and immunodeficiency disorders (Aisenberg. A.C. et al. (1985) N. Engl. J. Med. 313:529-533: Weiss. supra; and Olive. supra). Immunizations with peptides derived from TCRs are effective treatment for some human T-cell-mediated autoimmune disease and in animal models of such illnesses. in particular. rheumatoid arthritis (Bridges. S.L. and Moreland. L.W. (1998) Rheum. Dis. Clin. North Am. 24:641-650).

The discovery of new human cell surface receptor proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis. prevention. and treatment of cell proliferative disorders. immune system disorders. infections. and neuronal disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides. human cell surface receptor proteins. referred to collectively as "HCSRP" and individually as "HCSRP-1," "HCSRP-2," "HCSRP-3," "HCSRP-4," "HCSRP-5," "HCSRP-6," "HCSRP-7," "HCSRP-8," "HCSRP-9," "HCSRP-10," "HCSRP-11," "HCSRP-12," and "HCSRP-13." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13 and fragments thereof. The invention also includes a polypeptide comprising an amino acid sequence that differs by one or more conservative amino acid substitutions from an amino acid sequence selected from the group consisting of SEQ ID NO:1-13.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-13 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence

selected from the group consisting of SEQ ID NO:1-13 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex: and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:14-26 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:14-26 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:14-26 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-13. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing a polynucleotide of the invention under conditions suitable for the expression of the polypeptide: and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-13 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-13 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of HCSRP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-13 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of HCSRP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-13 and fragments thereof.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HCSRP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of HCSRP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding HCSRP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze HCSRP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be

used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"HCSRP" refers to the amino acid sequences of substantially purified HCSRP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of HCSRP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HCSRP either by directly interacting with HCSRP or by acting on components of the biological pathway in which HCSRP participates.

An "allelic variant" is an alternative form of the gene encoding HCSRP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HCSRP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as HCSRP or a polypeptide with at least one functional characteristic of HCSRP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HCSRP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HCSRP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HCSRP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HCSRP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine.

Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of HCSRP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HCSRP either by directly interacting with HCSRP or by acting on components of the biological pathway in which HCSRP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind HCSRP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the

complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HCSRP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HCSRP or fragments of HCSRP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded

as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of HCSRP or the polynucleotide encoding HCSRP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues

in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:14-26 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:14-26, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:14-26 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:14-26 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:14-26 and the region of SEQ ID NO:14-26 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-13 is encoded by a fragment of SEQ ID NO:14-26. A fragment of SEQ ID NO:1-13 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-13. For example, a fragment of SEQ ID NO:1-13 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-13. The precise length of a fragment of SEQ ID NO:1-13 and the region of SEQ ID NO:1-13 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention

include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of HCSRP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HCSRP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably

linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding HCSRP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the

PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HCSRP, or fragments thereof, or HCSRP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a

protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding

polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human cell surface receptor proteins (HCSRP), the polynucleotides encoding HCSRP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative disorders, immune system disorders, infections, and neuronal disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HCSRP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HCSRP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each HCSRP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods used to characterize each polypeptide through sequence homology and protein motifs. In

particular, the amino acid sequence of SEQ ID NO:1 from about amino acid residue 30 to about 81 is distinct from the tethered ligand thrombin receptor agonist peptide of the N-terminus of the human thrombin receptor and the amino acid sequence of SEQ ID NO:2 from about amino acid residue 115 to about 140 is distinct from the C-terminus joining and constant regions of the human TCRα subunit.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HCSRP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:14-26 and to distinguish between SEQ ID NO:14-26 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express HCSRP as a fraction of total tissues expressing HCSRP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing HCSRP as a fraction of total tissues expressing HCSRP. Column 5 lists the vectors used to subclone each cDNA library. Of particular note is the expression of HCSRP in cancer, autoimmune and inflammatory response, and in lung, thymus, bladder, seminal vesicle, and penile tissues, and in rheumatoid arthritis. In addition, SEQ ID NO: 14 is expressed primarily in tumor-associated epithelial tissues and SEQ ID NO: 15 is expressed primarily in growth- and tumor-associated epithelial tissues and in immune response tissues.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding HCSRP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses HCSRP variants. A preferred HCSRP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the HCSRP amino acid sequence, and which contains at least one functional or structural characteristic of HCSRP.

The invention also encompasses polynucleotides which encode HCSRP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:14-26, which encodes HCSRP.

The invention also encompasses a variant of a polynucleotide sequence encoding HCSRP. In particular, such a variant polynucleotide sequence will have at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HCSRP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID

NO:14-26 which has at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:14-26. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HCSRP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HCSRP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HCSRP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HCSRP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring HCSRP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HCSRP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HCSRP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HCSRP and HCSRP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HCSRP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:14-26 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of

the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HCSRP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HCSRP may be cloned in recombinant DNA molecules that direct expression of HCSRP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HCSRP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HCSRP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HCSRP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, HCSRP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HCSRP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other

proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active HCSRP, the nucleotide sequences encoding HCSRP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HCSRP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HCSRP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HCSRP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HCSRP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HCSRP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or

tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HCSRP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HCSRP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HCSRP into the vector's multiple cloning site disrupts the *lac*Z gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HCSRP are needed, e.g. for the production of antibodies, vectors which direct high level expression of HCSRP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HCSRP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of HCSRP. Transcription of sequences encoding HCSRP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HCSRP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain

infective virus which expresses HCSRP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HCSRP in cell lines is preferred. For example, sequences encoding HCSRP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is

also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HCSRP is inserted within a marker gene sequence, transformed cells containing sequences encoding HCSRP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HCSRP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HCSRP and that express HCSRP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HCSRP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HCSRP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HCSRP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HCSRP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HCSRP may be cultured under

conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HCSRP may be designed to contain signal sequences which direct secretion of HCSRP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HCSRP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HCSRP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HCSRP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HCSRP encoding sequence and the heterologous protein sequence, so that HCSRP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HCSRP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These

systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

Fragments of HCSRP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of HCSRP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HCSRP and human cell surface receptor proteins. In addition, the expression of HCSRP is closely associated with lung, thymus, bladder, seminal vesicle, and penile tissues, with rheumatoid arthritis, and with inflammation, cancer, and the nervous system. Therefore, HCSRP appears to play a role in cell proliferative disorders, immune system disorders, infections, and neuronal disorders. In the treatment of disorders associated with increased HCSRP expression or activity, it is desirable to decrease the expression or activity of HCSRP. In the treatment of disorders associated with decreased HCSRP expression or activity, it is desirable to increase the expression or activity of HCSRP.

Therefore, in one embodiment, HCSRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HCSRP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune system disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome,

episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; an infection caused by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, or togavirus; an infection caused by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, or campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, or mycoplasma; an infection caused by a fungal agent classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malasezzia, histoplasma, or other fungal agents causing various mycoses; an infection caused by a parasite classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas, tissue nematodes such as trichinella, intestinal nematodes such as ascaris, lymphatic filarial nematodes, trematodes such as schistosoma, or cestrodes such as tapeworm; and a neuronal disorder such as akathesia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar disorder, catatonia, cerebral neoplasms, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's disease, paranoid psychoses, postherpetic neuralgia, schizophrenia, and Tourette's disorder.

In another embodiment, a vector capable of expressing HCSRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HCSRP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HCSRP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HCSRP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HCSRP may be

administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HCSRP including, but not limited to, those listed above.

In a further embodiment, an antagonist of HCSRP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HCSRP. Examples of such disorders include, but are not limited to, those cell proliferative disorders, immune system disorders, infections, and neuronal disorders described above. In one aspect, an antibody which specifically binds HCSRP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express HCSRP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HCSRP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HCSRP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HCSRP may be produced using methods which are generally known in the art. In particular, purified HCSRP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HCSRP. Antibodies to HCSRP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HCSRP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HCSRP have an amino acid sequence consisting of at least about 5 amino acids, and generally will

consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HCSRP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HCSRP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HCSRP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HCSRP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HCSRP and its

specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HCSRP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HCSRP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of HCSRP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HCSRP epitopes, represents the average affinity, or avidity, of the antibodies for HCSRP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HCSRP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the HCSRP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HCSRP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of HCSRP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HCSRP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HCSRP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HCSRP. Thus, complementary molecules or fragments may be used to modulate HCSRP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HCSRP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HCSRP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding HCSRP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HCSRP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HCSRP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HCSRP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary

oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HCSRP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HCSRP, antibodies to HCSRP, and mimetics, agonists, antagonists, or inhibitors of HCSRP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any

number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's

solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HCSRP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HCSRP or fragments thereof, antibodies of HCSRP, and agonists, antagonists or inhibitors of HCSRP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such

as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HCSRP may be used for the diagnosis of disorders characterized by expression of HCSRP, or in assays to monitor patients being treated with HCSRP or agonists, antagonists, or inhibitors of HCSRP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HCSRP include methods which utilize the antibody and a label to detect HCSRP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HCSRP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HCSRP expression.

Normal or standard values for HCSRP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to HCSRP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of HCSRP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HCSRP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of HCSRP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HCSRP, and to monitor regulation of HCSRP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HCSRP or closely related molecules may be used to identify nucleic acid sequences which encode HCSRP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding HCSRP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the HCSRP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:14-26 or from genomic sequences including promoters, enhancers, and introns of the HCSRP gene.

Means for producing specific hybridization probes for DNAs encoding HCSRP include the cloning of polynucleotide sequences encoding HCSRP or HCSRP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HCSRP may be used for the diagnosis of disorders associated with expression of HCSRP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal

hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune system disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; an infection caused by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, or togavirus; an infection caused by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, or campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, or mycoplasma; an infection caused by a fungal agent classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malasezzia, histoplasma, or other fungal agents causing various mycoses; an infection caused by a parasite classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas, tissue nematodes such as trichinella, intestinal nematodes such as ascaris, lymphatic filarial nematodes, trematodes such as schistosoma, or cestrodes such as tapeworm; and a neuronal disorder such as akathesia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar disorder,

catatonia, cerebral neoplasms, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's disease, paranoid psychoses, postherpetic neuralgia, schizophrenia, and Tourette's disorder. The polynucleotide sequences encoding HCSRP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HCSRP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HCSRP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HCSRP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HCSRP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HCSRP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HCSRP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance

of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HCSRP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HCSRP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HCSRP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of HCSRP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HCSRP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price,

C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding HCSRP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HCSRP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HCSRP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HCSRP, or fragments thereof, and washed. Bound HCSRP is then detected by methods well known in the art. Purified HCSRP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HCSRP specifically compete with a test compound for binding HCSRP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HCSRP.

In additional embodiments, the nucleotide sequences which encode HCSRP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0636 P, filed November 12, 1998], U.S. Ser. No. [Attorney Docket No. PF-0650 P, filed December 7, 1998], and U.S. Ser. No. 60/123,404 are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA

purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared

using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS,

DOMO. PRODOM. Prosite. and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See. e.g.. Eddy. S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:14-26. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See. e.g., Sambrook. supra, ch. 7: Ausubel, 1995. supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition. the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{\% sequence identity} \times \text{\% maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40. the match will be exact within a 1% to 2% error. and. with a product score of 70. the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40. although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HCSRP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular. dermatologic. developmental. endocrine. gastrointestinal. hematopoietic/immune. musculoskeletal. nervous. reproductive. and urologic. The disease/condition categories included cancer. inflammation. trauma. cell proliferation. neurological. and pooled. For each category. the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of HCSRP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:14-26 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones

were religated using T4 ligase (New England Biolabs. Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:14-26 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:14-26 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and

compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the HCSRP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HCSRP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HCSRP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HCSRP-encoding transcript.

IX. Expression of HCSRP

Expression and purification of HCSRP is achieved using bacterial or virus-based expression systems. For expression of HCSRP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA

transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HCSRP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HCSRP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HCSRP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HCSRP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HCSRP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified HCSRP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of HCSRP Activity

An assay for HCSRP activity measures the expression of HCSRP on the cell surface. cDNA encoding HCSRP is transfected into an appropriate mammalian cell line. Cell surface proteins are labeled with biotin as described (de la Fuente, M. A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using HCSRP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of HCSRP expressed on the cell surface.

An alternative assay for HCSRP activity is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the amount of newly synthesized DNA in Swiss mouse 3T3 cells expressing HCSRP. An appropriate mammalian expression vector containing cDNA encoding HCSRP is added to quiescent 3T3 cultured cells using transfection methods well known in the art. The transfected cells are incubated in the presence of [^{3}H]thymidine and varying amounts of HCSRP ligand. Incorporation of [^{3}H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a tritium radioisotope counter, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold HCSRP ligand concentration range is indicative of receptor activity. One unit of activity per milliliter is defined as the concentration of HCSRP producing a 50% response level, where 100% represents maximal incorporation of [^{3}H]thymidine into acid-precipitable DNA. (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY, p. 73.)

XI. Functional Assays

HCSRP function is assessed by expressing the sequences encoding HCSRP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in

flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HCSRP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HCSRP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HCSRP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of HCSRP Specific Antibodies

HCSRP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HCSRP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-HCSRP activity by, for example, binding the peptide or HCSRP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring HCSRP Using Specific Antibodies

Naturally occurring or recombinant HCSRP is substantially purified by immunoaffinity chromatography using antibodies specific for HCSRP. An immunoaffinity column is constructed by covalently coupling anti-HCSRP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HCSRP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HCSRP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt

antibody/HCSRP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HCSRP is collected.

XIV. Identification of Molecules Which Interact with HCSRP

HCSRP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HCSRP, washed, and any wells with labeled HCSRP complex are assayed. Data obtained using different concentrations of HCSRP are used to calculate values for the number, affinity, and association of HCSRP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	14	2355971	LUNGNOT20	2355971F6 and 2355971H1 (LUNGNOT20)
2	15	2917059	THYMFET03	2917059F6, 2917059H1, and 2917059T6 (THYMFET03)
3	16	538181	LNODNOT02	538181H1 (LNODNOT02), 1437379F1 (PANCNOT08), 2945139F6 (BRAITUT23), SAJA00666R1, SAJA00920R1, SAJA00173F1, SAJA00783R1
4	17	1368760	SCORNON02	1368760H1 and 1368760T6 (SCORNON02), SAUB01170F1
5	18	1670669	BMARNOT03	1670669F6 and 1670669H1 (BMARNOT03), SAEA03216R1
6	19	2851578	BRSTTUT13	992052T1 (COLNNOT11), 1579409F1 (DUODNOT01), 1880543T6 (LEUKNOT03), 4088508H1 (LIVRNOT06), 2851578H1, 2851578X310F1, and 2851578X311F1 (BRSTTUT13)
7	20	3393757	LUNGNOT28	318405H1 (EOSIHET02), 2230506H1 (PROSNOT16), 4592475F6 (MASTTXT01), 3393757H1, 3393757X301D2, and 3393757X305B2 (LUNGNOT28), SZAH04055F1, SZAH01718F1
8	21	312256	LUNGNOT02	306245T6 (HEARNOT01), 312256H1 (LUNGNOT02), 4175078H1 (SINTNOT21), SCAA06277V1, SCAA02060V1, SCAA02643V1
9	22	1615704	BRAITUT12	1615704F6 (BRAITUT12), 1615704H1 (BRAITUT12), 1615704X13 (BRAITUT12), 1727857F6 (PROSNOT14), 3539783H1 (SEMVNOT04)
10	23	1659465	URETTUT01	265886F1 (HNT2AGT01), 1441970R1 (THYRNOT03), 1522238X12 (BLADTUT04), 1659465H1 (URETTUT01), 2187904T6 (PROSNOT26), 3237635H1 (COLAUCT01)

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
11	24	2120743	BRSTNOT07	661461R6 (BRAINOT03), 1318546F6 (BLADNOT04), 1318694T1 (BLADNOT04), 2044109H1 (HIPONON02), 2120743F6 (BRSTNOT07), 2120743H1 (BRSTNOT07), SBIA06480D1, SBIA02214D1, SBIA05533D1, SBIA03320D1, SBIA10018D1, SBIA07320D1, SBIA00175D1
12	25	3344986	SPLNNOT09	138687F1 (LIVRNOT01), 138687X27C1 (LIVRNOT01), 169384X8 (BMARNOR02), 3344986H1 (SPLNNOT09)
13	26	3576503	BRONNOT01	2525456F6 (BRAITUT21), 2527260F6 (BRAITUT21), 3576503H1 (BRONNOT01), 3742127F6 (MENTNOT01)

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Motifs, Signature Sequences, and Protein Domains	Homologous Sequences	Analytical Methods
1	81	S21 T43		Signal peptide: M1-A22 Receptor signature: C12-R28	thrombin receptor g339677	BLAST MOTIFS BLOCKS PRINTS HMM
2	140	S20 S33 S49 S73 S61 T90	N88	Signal peptide: M1-G22 Receptor tyrosine kinase signature: R82-A99 Receptor signature: A99-C111	T cell receptor alpha chain g1223888	BLAST MOTIFS BLOCKS PRINTS HMM
3	358	T47 T65 T163 T228 T283 S299 T352		Ig domains: G91-G144; G184-A241 Signal peptide: M1-A25	IgG Fc receptor g583604	BLAST MOTIFS PFAM SPSCAN HMM

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Motifs, Signature Sequences, and Protein Domains	Homologous Sequences	Analytical Methods
4	201	T44 T131 T65 T167	N63	Glutamate receptor domain: M1-G170 GPCR/Neurotransmitter receptor signatures: T114-T122; M7-S26 L145-H155; F90-G103 Signal peptide: M1-G28	Glutamate receptor isoform g507827	BLAST MOTIFS PFAM PRINTS SPSCAN
5	117	T54 T75 S104		Cytokine receptor signature: D32-E93 Signal peptide: M1-W24 or M1-A25	EPO receptor isoform g553281	BLAST MOTIFS PROFILESCAN SPSCAN HMM
6	455	S22 S380 S53 T189 T206 S271 S281 T310 S430	N76 N207	Rhodopsin GPCR domain: G141-Y387 Rhodopsin signatures: F126-C150; N159-R180 L204-V226; A325-A349 F369-L395 Transmembrane domains: L291-T310; A128-C150	Rhodopsin-like GPCR/HM74 g219867	BLAST MOTIFS PROFILESCAN PFAM BLOCKS PRINTS HMM

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Motifs, Signature Sequences, and Protein Domains	Homologous Sequences	Analytical Methods
7	453	S71 T418 S87 T244 T340 S428 T433 T438 T449	N73 N77 N183 N247 N252	Secretin signatures: V155-K179; I218-L241 K261-L286; W303-K328 A377-L398 Transmembrane domains: M159-L177; I302-V322 Y378-L398	Secretin-like GPCR/EMR1 g784994	BLAST MOTIFS BLOCKS PRINTS HMM
8	442	T53 S103 S115 T155 S190 T176 T233 T310 S368 Y241 Y329	N67 N101 N113 N165 N304 N308 N432	Transmembrane domain: A375-I394 Signal peptide: M1-G44 Immunoglobulin domains: G57-L126; G159-V222 G260-A315	Immuno-superfamily protein B12 g3779242	HMM BLAST PFAM SPSCAN BLOCKS-DOMO
9	382	S152 S199 T267 T324 S29	N65 N178 N322	Transmembrane domains: M1-I21; L289-F306 Signal peptide M1-S29 Somatostatin receptor: L237-L254 2 Poly-Immunoglobulin receptor: F9-P46; E86-P123 P104-F134; V259-F306 L283-A320		HMM SPSCAN PRINTS BLOCKS-DOMO

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Motifs, Signature Sequences, and Protein Domains	Homologous Sequences	Analytical Methods
10	257	S162 S7 T228		Transmembrane domain: Y187-L210 Signal peptide: M1-S33 Endothelin-A receptor: Q236-T255 2 Poly-Immunoglobulin receptor: Y91-R138; Y170-L217		HMM SPSCAN PRINTS BLOCKS-DOMO
11	697	S266 S623 T18 S138 T155 S215 S247 T524 T631 T684 S4 S32 S113 S122 T219 T249 T332 T519 T606 T611 S619	N111 N533 N598	Transmembrane domains: L36-V54; S62-Y84 Y125-R142; S177-I192 S198-F214; F298-H320 I381-G398 Signal peptide: M1-W56 Prostanoid EP1 receptor: V360-S384; D110-R134 2 Poly-Immunoglobulin receptor: I64-N111; L162-A209		HMM SPSCAN PRINTS BLOCKS-DOMO

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Motifs, Signature Sequences, and Protein Domains	Homologous Sequences	Analytical Methods
12	325	S56 T185 T234 S289 S76 S211 T245 S248	N64 N287	Transmembrane domain: Q26-V44 Signal peptide: M1-S52 2 Poly-Immunoglobulin receptor: M1-P48; A22-K69 V95-G142; V118-G165	Non-CD4 glycoprotein gp120 receptor R32188	HMM BLAST SPSCAN BLOCKS-DOMO
13	369	T12 T69 T299 S312 S330 T346 S50 S92 T145	N10 N90 N297	Transmembrane domains: I144-L161; V196-L219 L264-Y283 Signal peptide: M1-G47 Prostaglandin D receptor: Y6-A20; L41-P58 L93-L104; I185-S200 P236-E259; P280-K291 2 Poly-Immunoglobulin receptor: T11-P58; N10-L57 R52-A99; S195-L225 7 Transmembrane receptor: G33-P280	DP prostanoid receptor g940379	HMM BLAST SPSCAN PRINTS BLOCKS-DOMO PFAM

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Diseases, Disorders, or Conditions (Fraction of Total)	Vector
14	146-301	Lung (1.000)	Cancer (1.000)	pINCY
15	423-500	Fetal Thymus (0.333) Wrist Synovium (0.333) Seminal Vesicle (0.333)	Cancer (0.333) Cell Proliferation (0.333) Inflammation (0.333)	pINCY
16	94-138	Hematopoietic/Immune (0.636) Gastrointestinal (0.273)	Inflammation (0.636) Cancer (0.364)	PSPORT1
17	703-747	Nervous System (1.000)	Neurological (0.333) Trauma (0.333) Cancer (0.333)	PSPORT1
18	293-325 434-478	Cardiovascular (0.333) Developmental (0.167) Hematopoietic/Immune (0.167)	Cancer (0.833) Fetal (0.167)	pINCY
19	109-153 1339-1383	Gastrointestinal (0.333) Reproductive (0.333) Hematopoietic/Immune (0.250)	Cancer (0.500) Inflammation (0.250) Fetal (0.167)	pINCY
20	519-563	Hematopoietic/Immune (0.571) Cardiovascular (0.143) Reproductive (0.143)	Inflammation (0.357) Cancer (0.214) Trauma (0.143)	pINCY
21	319-360	Nervous (0.235) Cardiovascular (0.206) Reproductive (0.206)	Cancer and Cell Proliferation (0.529) Inflammation (0.412)	PBLUESCRIPT
22	558-599	Reproductive (0.361) Nervous (0.180) Cardiovascular (0.115)	Cancer and Cell Proliferation (0.689) Inflammation (0.311)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Diseases, Disorders, or Conditions (Fraction of Total)	Vector
23	8-47 900-940	Reproductive (0.280) Nervous (0.160) Gastrointestinal (0.133)	Cancer and Cell Proliferation (0.666) Inflammation (0.253)	pINCY
24	800-840	Reproductive (0.237) Nervous (0.186) Cardiovascular (0.153)	Cancer and Cell Proliferation (0.664) Trauma (0.186)	pINCY
25	770-814	Hematopoietic/Immune (0.333) Gastrointestinal (0.222) Cardiovascular (0.111) Endocrine (0.111) Nervous (0.111) Reproductive (0.111)	Inflammation (0.444) Fetal (0.222) Trauma (0.222)	pINCY
26	1078-1119	Nervous (0.500) Cardiovascular (0.250) Musculoskeletal (0.250)	Cancer (0.750) Trauma (0.250)	pINCY

Table 4

Polypeptide SEQ ID NO:	Library	Library Description
14	LUNGNOT20	This library was constructed using RNA isolated from right upper lobe lung tissue removed from a 61-year-old Caucasian male. Pathology indicated panacinal emphysema with blebs in the right anterior upper lobe and apex, as well as emphysema in the right posterior upper lobe. Patient history included angina pectoris and gastric ulcer. Family history included a subdural hemorrhage, cancer, atherosclerotic coronary artery disease, and pneumonia.
15	THYMFET03	This library was constructed using RNA isolated from thymus tissue removed from a premature Caucasian male fetus who died at birth.
16	LNODNOT02	This library was constructed using RNA isolated from the lymph node tissue of a 42-year-old Caucasian female who died of cardiac arrest.
17	SCORNON02	This normalized spinal cord library was constructed from 3.24 million independent clones from a spinal cord library. RNA was isolated from spinal cord tissue removed from a 71-year-old Caucasian male who died from respiratory arrest. Patient history included myocardial infarction, gangrene, and end stage renal disease. The normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9928).
18	BMARNOT03	This library was constructed using RNA isolated from left tibial bone marrow tissue of a 16-year-old Caucasian male during a partial left tibial ostectomy with free skin graft. Patient history included an abnormality of the red blood cells. Previous surgeries included bone and bone marrow biopsy and soft tissue excision. Family history included osteoarthritis.
19	BRSTTUT13	This library was constructed using RNA isolated from breast tumor tissue removed from a 46-year-old Caucasian female during a unilateral extended simple mastectomy with breast reconstruction. Pathology indicated an invasive grade 3 adenocarcinoma, ductal type with apocrine features and greater than 50% intraductal component. Patient history included breast cancer.

Table 4 (cont.)

Polypeptide SEQ ID NO:	Library	Library Description
20	LUNGNOT28	This library was constructed using RNA isolated from lung tissue removed from a 53-year-old male. Pathology for the associated tumor tissue indicated grade 4 adenocarcinoma.
21	LUNGNOT02	This library was constructed using RNA isolated from the lung tissue of a 47-year-old Caucasian male, who died of a subarachnoid hemorrhage.
22	BRAITUT12	This library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated grade 4 gemistocytic astrocytoma.
23	URETTUT01	This library was constructed using RNA isolated from right ureter tumor tissue of a 69-year-old Caucasian male during ureterectomy and lymph node excision. Pathology indicated invasive grade 3 transitional cell carcinoma. Patient history included benign colon neoplasm, asthma, emphysema, acute duodenal ulcer, and hyperplasia of the prostate. Family history included atherosclerotic coronary artery disease, congestive heart failure, and malignant lung neoplasm.
24	BRSTNOT07	This library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.
25	SPLNNOT09	This library was constructed using RNA isolated from diseased spleen tissue removed from a 22-year-old Caucasian male during a total splenectomy. Pathology indicated Gaucher's disease with marked splenomegaly. Patient history included thyroid disorders and type 1 Gaucher's disease.

Table 4 (cont.)

Polypeptide SEQ ID NO:	Library	Library Description
26	BRONNOT01	This library was constructed using RNA isolated from bronchial tissue removed from a 15-year-old Caucasian male.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥ GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:14-26 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide

sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of HCSRP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of HCSRP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
TANG, Y. Tom
CORLEY, Neil C.
GUEGLER, Karl J.
YUE, Henry
BAUGHN, Mariah R.
LAL, Preeti
HILLMAN, Jennifer L.
BANDMAN, Olga
AZIMZAI, Yalda
AU-YOUNG, Janice

<120> HUMAN CELL SURFACE RECEPTOR PROTEINS

<130> PF-0636 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/121,280; unassigned; 09/206,647; unassigned; 60/123,404
<151> 1998-11-12; 1998-11-12; 1998-12-07; 1998-12-07; 1999-03-08

<160> 26

<170> PERL Program

<210> 1
<211> 81
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2355971CD1

<400> 1
Met Gly Pro Arg Arg Leu Leu Leu Val Ala Ala Cys Phe Ser Leu
1 5 10 15
Cys Gly Pro Leu Leu Ser Ala Arg Thr Arg Ala Arg Arg Pro Gly
20 25 30
Glu Arg Cys Thr Gly Met Gly Cys Ala Gly Gly Gly Thr Pro Arg
35 40 45
Gly Asp Cys Gly Gly His Cys Cys Asp Phe Ser Ser Pro Leu Pro
50 55 60
Gln Phe Pro Pro Lys Ala Lys Leu Ala Phe Gly Leu Arg Ser Gly
65 70 75
Val Phe Ser Ser His Val
80

<210> 2
<211> 140
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2917059CD1

<400> 2
Met Met Ala Gly Ile Arg Ala Leu Phe Met Tyr Leu Trp Leu Gln
1 5 10 15
Leu Asp Trp Val Ser Arg Gly Glu Ser Val Gly Leu His Leu Pro
20 25 30
Thr Leu Ser Val Gln Glu Gly Asp Asn Ser Ile Ile Asn Cys Ala
35 40 45
Tyr Ser Asn Ser Ala Ser Asp Tyr Phe Ile Trp Tyr Lys Gln Glu
50 55 60
Ser Gly Lys Gly Pro Gln Phe Ile Ile Asp Ile Arg Ser Asn Met
65 70 75
Asp Lys Arg Gln Gly Gln Arg Val Thr Val Leu Leu Asn Lys Thr
80 85 90
Val Lys His Leu Ser Leu Gln Ile Ala Ala Thr Gln Pro Gly Asp
95 100 105
Ser Ala Val Tyr Phe Cys Ala Glu Asn Thr His Cys Phe Pro Gly
110 115 120
Ile Cys Asn His His Pro Asn Leu Arg Trp Glu Val Lys Gln His
125 130 135
Pro Phe Pro Leu Gln
140

<210> 3
<211> 358
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 538181CD1

<400> 3
Met Ala Trp Ala Leu Tyr Leu Ser Leu Gly Val Leu Trp Val Ala
1 5 10 15
Gln Met Leu Leu Ala Ala Gly Cys His Ala Ala Ala Ser Phe Glu
20 25 30
Thr Leu Gln Cys Glu Gly Pro Val Cys Thr Glu Glu Ser Ser Cys
35 40 45
His Thr Glu Asp Asp Leu Thr Asp Ala Arg Glu Ala Gly Phe Gln
50 55 60
Val Lys Ala Tyr Thr Phe Ser Glu Pro Phe His Leu Ile Val Ser
65 70 75
Tyr Asp Trp Leu Ile Leu Gln Gly Pro Ala Lys Pro Val Phe Glu
80 85 90
Gly Asp Leu Leu Val Leu Arg Cys Gln Ala Trp Gln Asp Trp Pro
95 100 105
Leu Thr Gln Val Thr Phe Tyr Arg Asp Gly Ser Ala Leu Gly Pro
110 115 120

```
Pro Gly Pro Asn Arg Glu Phe Ser Ile Thr Val Val Gln Lys Ala
                125                 130                 135
Asp Ser Gly His Tyr His Cys Ser Gly Ile Phe Gln Ser Pro Gly
                140                 145                 150
Pro Gly Ile Pro Glu Thr Ala Ser Val Val Ala Ile Thr Val Gln
                155                 160                 165
Glu Leu Phe Pro Ala Pro Ile Leu Arg Ala Leu Pro Ser Ala Glu
                170                 175                 180
Pro Gln Ala Gly Gly Pro Met Thr Leu Ser Cys Gln Thr Lys Leu
                185                 190                 195
Pro Leu Gln Arg Ser Ala Ala Arg Leu Leu Phe Ser Phe Tyr Lys
                200                 205                 210
Asp Gly Arg Ile Val Gln Ser Arg Gly Leu Ser Ser Glu Phe Gln
                215                 220                 225
Ile Pro Thr Ala Ser Glu Asp His Ser Gly Ser Tyr Trp Cys Glu
                230                 235                 240
Ala Ala Thr Glu Asp Asn Gln Val Trp Lys Gln Ser Pro Gln Leu
                245                 250                 255
Glu Ile Arg Val Gln Gly Ala Ser Ser Ser Ala Ala Pro Pro Thr
                260                 265                 270
Leu Asn Pro Ala Pro Gln Lys Ser Ala Ala Pro Gly Thr Ala Pro
                275                 280                 285
Glu Glu Ala Pro Gly Pro Leu Pro Pro Pro Pro Thr Pro Ser Ser
                290                 295                 300
Glu Asp Pro Gly Phe Ser Ser Pro Leu Gly Met Pro Asp Pro His
                305                 310                 315
Leu Tyr His Gln Met Gly Leu Leu Leu Lys His Met Gln Asp Val
                320                 325                 330
Arg Val Leu Leu Gly His Leu Leu Met Glu Leu Arg Glu Leu Ser
                335                 340                 345
Gly His Arg Lys Pro Gly Thr Thr Lys Ala Thr Ala Glu
                350                 355
```

<210> 4
<211> 201
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1368760CD1

<400> 4

```
Met Ala Arg Gln Lys Lys Met Gly Gln Ser Val Leu Arg Ala Val
  1                 5                  10                  15
Phe Phe Leu Val Leu Gly Leu Leu Gly His Ser His Gly Gly Phe
                20                  25                  30
Pro Asn Thr Ile Ser Ile Gly Gly Leu Phe Met Arg Asn Thr Val
                35                  40                  45
Gln Glu His Ser Ala Phe Arg Phe Ala Val Gln Leu Tyr Asn Thr
                50                  55                  60
Asn Gln Asn Thr Thr Glu Lys Pro Phe His Leu Asn Tyr His Val
                65                  70                  75
Asp Leu Leu Asp Ser Ser Asn Ser Phe Ser Val Thr Asn Ala Phe
```

```
                  80                  85                  90
Cys Ser Gln Phe Ser Arg Gly Val Tyr Ala Ile Phe Gly Phe Tyr
                  95                 100                 105
Asp Gln Met Ser Met Asn Thr Leu Thr Ser Phe Cys Gly Ala Leu
                 110                 115                 120
His Thr Ser Phe Val Thr Pro Ser Phe Pro Thr Asp Ala Asp Val
                 125                 130                 135
Gln Phe Val Ile Gln Met Arg Pro Ala Leu Lys Gly Ala Ile Leu
                 140                 145                 150
Ser Leu Leu Gly His Tyr Lys Trp Glu Lys Phe Val Tyr Leu Tyr
                 155                 160                 165
Asp Thr Glu Arg Gly Lys Lys Arg His Leu Leu Cys Ser Leu Asp
                 170                 175                 180
Ile His Val Ile Val Phe Lys Leu Pro Gln Leu Met Cys Pro Leu
                 185                 190                 195
Leu Pro Ile Asn Lys Ile
                 200

<210> 5
<211> 117
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1670669CD1

<400> 5
Met Asp His Leu Gly Ala Ser Leu Trp Pro Gln Val Gly Ser Leu
  1               5                  10                  15
Cys Leu Leu Leu Ala Gly Ala Ala Trp Ala Pro Pro Pro Asn Leu
                 20                  25                  30
Pro Asp Pro Lys Phe Glu Ser Lys Ala Ala Leu Leu Ala Ala Arg
                 35                  40                  45
Gly Pro Glu Glu Leu Leu Cys Phe Thr Glu Arg Val Gly Gly Leu
                 50                  55                  60
Gly Met Ser His Gly Lys Leu Cys Arg Leu His Gln Ala Pro Thr
                 65                  70                  75
Ala Arg Gly Gly Gly Ala Leu Leu Val Cys Ala Ala Tyr Arg Arg
                 80                  85                  90
His Val Glu Leu Arg Ala Pro Arg Val Gly Arg His Ser Ser Leu
                 95                 100                 105
Arg Arg Ser Ala Ile Ser Pro Cys His Pro His Gln
                110                 115

<210> 6
<211> 455
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2851578CD1
```

<400> 6

```
Met Glu Gly Gly Arg Ala Ala Gly Phe Pro Leu Ala Cys Lys Gln
  1               5                  10                  15
Glu His Arg Val Phe Leu Ser Gly Cys Glu Asn Ala Asp Glu Asn
             20                  25                  30
Pro Arg Met Leu Cys His Arg Gly Gly Gln Leu Ile Val Pro Ile
             35                  40                  45
Ile Pro Leu Cys Pro Glu His Ser Cys Arg Gly Arg Arg Leu Gln
             50                  55                  60
Asn Leu Leu Ser Gly Pro Trp Pro Lys Gln Pro Met Glu Leu His
             65                  70                  75
Asn Leu Ser Ser Pro Ser Pro Ser Leu Ser Ser Ser Val Leu Pro
             80                  85                  90
Pro Ser Phe Ser Pro Ser Pro Ser Ser Ala Pro Ser Ala Phe Thr
             95                 100                 105
Thr Val Gly Gly Ser Ser Gly Gly Pro Cys His Pro Thr Ser Ser
            110                 115                 120
Ser Leu Val Ser Ala Phe Leu Ala Pro Ile Leu Ala Leu Glu Phe
            125                 130                 135
Val Leu Gly Leu Val Gly Asn Ser Leu Ala Leu Phe Ile Phe Cys
            140                 145                 150
Ile His Thr Arg Pro Trp Thr Ser Asn Thr Val Phe Leu Val Ser
            155                 160                 165
Leu Val Ala Ala Asp Phe Leu Leu Ile Ser Asn Leu Pro Leu Arg
            170                 175                 180
Val Asp Tyr Tyr Leu Leu His Glu Thr Trp Arg Phe Gly Ala Ala
            185                 190                 195
Ala Cys Lys Val Asn Leu Phe Met Leu Ser Thr Asn Arg Thr Ala
            200                 205                 210
Ser Val Val Phe Leu Thr Ala Ile Ala Leu Asn Arg Tyr Leu Lys
            215                 220                 225
Val Val Gln Pro His His Val Leu Ser Arg Ala Ser Val Gly Ala
            230                 235                 240
Ala Ala Arg Val Ala Gly Gly Leu Trp Val Gly Ile Leu Leu Leu
            245                 250                 255
Asn Gly His Leu Leu Leu Ser Thr Phe Ser Gly Pro Ser Cys Leu
            260                 265                 270
Ser Tyr Arg Val Gly Thr Lys Pro Ser Ala Ser Leu Arg Trp His
            275                 280                 285
Gln Ala Leu Tyr Leu Leu Glu Phe Phe Leu Pro Leu Ala Leu Ile
            290                 295                 300
Leu Phe Ala Ile Val Ser Ile Gly Leu Thr Ile Arg Asn Arg Gly
            305                 310                 315
Leu Gly Gly Gln Ala Gly Pro Gln Arg Ala Met Arg Val Leu Ala
            320                 325                 330
Met Val Val Ala Val Tyr Thr Ile Cys Phe Leu Pro Ser Ile Ile
            335                 340                 345
Phe Gly Met Ala Ser Met Val Ala Phe Trp Leu Ser Ala Cys Arg
            350                 355                 360
Ser Leu Asp Leu Cys Thr Gln Leu Phe His Gly Ser Leu Ala Phe
            365                 370                 375
Thr Tyr Leu Asn Ser Val Leu Asp Pro Val Leu Tyr Cys Phe Ser
            380                 385                 390
Ser Pro Asn Phe Leu His Gln Ser Arg Ala Leu Leu Gly Leu Thr
            395                 400                 405
```

```
Arg Gly Arg Gln Gly Pro Val Ser Asp Glu Ser Ser Tyr Gln Pro
                410                 415                 420
Ser Arg Gln Trp Arg Tyr Arg Glu Ala Ser Arg Lys Ala Glu Ala
                425                 430                 435
Ile Gly Lys Leu Lys Val Gln Gly Glu Val Ser Leu Glu Lys Glu
                440                 445                 450

Gly Ser Ser Gln Gly
                455


<210> 7
<211> 453
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3393757CD1

<400> 7
Met Gln Leu Asp Trp Asn Gln Ala Gln Lys Ser Gly Asp Pro Gly
  1               5                  10                  15
Pro Ser Val Val Gly Leu Val Ser Ile Pro Gly Met Gly Lys Leu
                20                  25                  30
Leu Ala Glu Ala Pro Leu Val Leu Glu Pro Glu Lys Gln Met Leu
                35                  40                  45
Leu His Glu Thr His Gln Gly Leu Leu Gln Asp Gly Ser Pro Ile
                50                  55                  60
Leu Leu Ser Asp Val Ile Ser Ala Phe Leu Ser Asn Asn Asp Thr
                65                  70                  75
Gln Asn Leu Ser Ser Pro Val Thr Phe Thr Phe Ser His Arg Ser
                80                  85                  90
Val Ile Pro Arg Gln Lys Val Leu Cys Val Phe Trp Glu His Gly
                95                 100                 105
Gln Asn Gly Cys Gly His Trp Ala Thr Thr Gly Cys Ser Thr Ile
                110                 115                 120
Gly Thr Arg Asp Thr Ser Thr Ile Cys Arg Cys Thr His Leu Ser
                125                 130                 135
Ser Phe Ala Val Leu Met Ala His Tyr Asp Val Gln Glu Glu Asp
                140                 145                 150
Pro Val Leu Thr Val Ile Thr Tyr Met Gly Leu Ser Val Ser Leu
                155                 160                 165
Leu Cys Leu Leu Leu Ala Ala Leu Thr Phe Leu Leu Cys Lys Ala
                170                 175                 180
Ile Gln Asn Thr Ser Thr Ser Leu His Leu Gln Leu Ser Leu Cys
                185                 190                 195
Leu Phe Leu Ala His Leu Leu Phe Leu Val Ala Ile Asp Gln Thr
                200                 205                 210
Gly His Lys Val Leu Cys Ser Ile Ile Ala Gly Thr Leu His Tyr
                215                 220                 225
Leu Tyr Leu Ala Thr Leu Thr Trp Met Leu Leu Glu Ala Leu Tyr
                230                 235                 240
Leu Phe Leu Thr Ala Arg Asn Leu Thr Val Val Asn Tyr Ser Ser
                245                 250                 255
```

```
Ile Asn Arg Phe Met Lys Lys Leu Met Phe Pro Val Gly Tyr Gly
                260                 265                 270
Val Pro Ala Val Thr Val Ala Ile Ser Ala Ala Ser Arg Pro His
                275                 280                 285
Leu Tyr Gly Thr Pro Ser Arg Cys Trp Leu Gln Pro Glu Lys Gly
                290                 295                 300
Phe Ile Trp Gly Phe Leu Gly Pro Val Cys Ala Ile Phe Ser Val
                305                 310                 315
Asn Leu Val Leu Phe Leu Val Thr Leu Trp Ile Leu Lys Asn Arg
                320                 325                 330
Leu Ser Ser Leu Asn Ser Glu Val Ser Thr Leu Arg Asn Thr Arg
                335                 340                 345
Met Leu Ala Phe Lys Ala Thr Ala Gln Leu Phe Ile Leu Gly Cys
                350                 355                 360
Thr Trp Cys Leu Gly Ile Leu Gln Val Gly Pro Ala Ala Arg Val
                365                 370                 375
Met Ala Tyr Leu Phe Thr Ile Ile Asn Ser Leu Gln Gly Val Phe
                380                 385                 390
Ile Phe Leu Val Tyr Cys Leu Leu Ser Gln Gln Val Arg Glu Gln
                395                 400                 405
Tyr Gly Lys Trp Ser Lys Gly Ile Arg Lys Leu Lys Thr Glu Ser
                410                 415                 420
Glu Met His Thr Leu Ser Ser Ser Ala Lys Ala Asp Thr Ser Lys
                425                 430                 435
Pro Ser Thr Val Arg Ser Arg Ile Ala Pro Glu His Phe Thr Asn
                440                 445                 450
Arg Pro Thr
```

<210> 8
<211> 442
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 312256CD1

<400> 8

```
Met Ala Ser Val Val Leu Pro Ser Gly Ser Gln Cys Ala Ala Ala
  1               5                  10                  15
Ala Ala Ala Ala Ala Pro Pro Gly Leu Arg Leu Arg Leu Leu Leu
                20                  25                  30
Leu Leu Phe Ser Ala Ala Ala Leu Ile Pro Thr Gly Asp Gly Gln
                35                  40                  45
Asn Leu Phe Thr Lys Asp Val Thr Val Ile Glu Gly Glu Val Ala
                50                  55                  60
Thr Ile Ser Cys Gln Val Asn Lys Ser Asp Asp Ser Val Ile Gln
                65                  70                  75
Leu Leu Asn Pro Asn Arg Gln Thr Ile Tyr Phe Arg Asp Phe Arg
                80                  85                  90
Pro Leu Lys Asp Ser Arg Phe Gln Leu Leu Asn Phe Ser Ser Ser
                95                 100                 105
Glu Leu Lys Val Ser Leu Thr Asn Val Ser Ile Ser Asp Glu Gly
```

```
                110                 115                 120
Arg Tyr Phe Cys Gln Leu Tyr Thr Asp Pro Pro Gln Glu Ser Tyr
                125                 130                 135
Thr Thr Ile Thr Val Leu Val Pro Pro Arg Asn Leu Met Ile Asp
                140                 145                 150
Ile Gln Lys Asp Thr Ala Val Glu Gly Glu Glu Ile Glu Val Asn
                155                 160                 165
Cys Thr Ala Met Ala Ser Lys Pro Ala Thr Thr Ile Arg Trp Phe
                170                 175                 180
Lys Gly Asn Thr Glu Leu Lys Gly Lys Ser Glu Val Glu Glu Trp
                185                 190                 195
Ser Asp Met Tyr Thr Val Thr Ser Gln Leu Met Leu Lys Val His
                200                 205                 210
Lys Glu Asp Asp Gly Val Pro Val Ile Cys Gln Val Glu His Pro
                215                 220                 225
Ala Val Thr Gly Asn Leu Gln Thr Gln Arg Tyr Leu Glu Val Gln
                230                 235                 240
Tyr Lys Pro Gln Val His Ile Gln Met Thr Tyr Pro Leu Gln Gly
                245                 250                 255
Leu Thr Arg Glu Gly Asp Ala Leu Glu Leu Thr Cys Glu Ala Ile
                260                 265                 270
Gly Lys Pro Gln Pro Val Met Val Thr Trp Val Arg Val Asp Asp
                275                 280                 285
Glu Met Pro Gln His Ala Val Leu Ser Gly Pro Asn Leu Phe Ile
                290                 295                 300
Asn Asn Leu Asn Lys Thr Asp Asn Gly Thr Tyr Arg Cys Glu Ala
                305                 310                 315
Ser Asn Ile Val Gly Lys Ala His Ser Asp Tyr Met Leu Tyr Val
                320                 325                 330
Tyr Asp Pro Pro Thr Thr Ile Pro Pro Pro Thr Thr Thr Thr Thr
                335                 340                 345
Thr Thr Thr Thr Thr Thr Thr Thr Ile Leu Thr Ile Ile Thr Asp
                350                 355                 360
Ser Arg Ala Gly Glu Glu Gly Ser Ile Arg Ala Val Asp His Ala
                365                 370                 375
Val Ile Gly Gly Val Val Ala Val Val Val Phe Ala Met Leu Cys
                380                 385                 390
Leu Leu Ile Ile Leu Gly Arg Tyr Phe Ala Arg His Lys Gly Thr
                395                 400                 405
Tyr Phe Thr His Glu Ala Lys Gly Ala Asp Asp Ala Ala Asp Ala
                410                 415                 420
Asp Thr Ala Ile Ile Asn Ala Glu Gly Gly Gln Asn Asn Ser Glu
                425                 430                 435
Glu Lys Lys Glu Tyr Phe Ile
                440
```

<210> 9
<211> 382
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1615704CD1

<400> 9

```
Met Asp Phe Leu Val Leu Phe Leu Phe Tyr Leu Ala Ser Val Leu
 1               5                  10                  15
Met Gly Leu Val Leu Ile Cys Val Cys Ser Lys Thr His Ser Leu
            20                  25                  30
Lys Gly Leu Ala Arg Gly Gly Ala Gln Ile Phe Ser Cys Ile Ile
        35                  40                  45
Pro Glu Cys Leu Gln Arg Ala Val His Gly Leu Leu His Tyr Leu
    50                  55                  60
Phe His Thr Arg Asn His Thr Phe Ile Val Leu His Leu Val Leu
65                  70                  75
Gln Gly Met Val Tyr Thr Glu Tyr Thr Trp Glu Val Phe Gly Tyr
                80                  85                  90
Cys Gln Glu Leu Glu Leu Ser Leu His Tyr Leu Leu Leu Pro Tyr
            95                  100                 105
Leu Leu Leu Gly Val Asn Leu Phe Phe Phe Thr Leu Thr Cys Gly
        110                 115                 120
Thr Asn Pro Gly Ile Ile Thr Lys Ala Asn Glu Leu Leu Phe Leu
    125                 130                 135
His Val Tyr Glu Phe Asp Glu Val Met Phe Pro Lys Asn Val Arg
140                 145                 150
Cys Ser Thr Cys Asp Leu Arg Lys Pro Ala Arg Ser Lys His Cys
155                 160                 165
Ser Glu Cys Gly Ser Arg Asp Ser Ser Gly Thr Ser Asn Ser Thr
                170                 175                 180
Cys Val Gly Phe Val Cys Glu Gly Met Phe Pro Glu Ser Glu Ser
            185                 190                 195
Arg Ala Ser Ser Pro Pro Asp Met Val Cys Val Thr Trp Cys Val
        200                 205                 210
His Arg Phe Asp His His Cys Val Trp Val Asn Asn Cys Ile Gly
    215                 220                 225
Ala Trp Asn Ile Arg Tyr Phe Leu Ile Tyr Val Leu Thr Leu Thr
230                 235                 240
Ala Ser Ala Ala Thr Val Ala Ile Val Ser Thr Thr Phe Leu Val
                245                 250                 255
His Leu Val Val Met Ser Asp Leu Tyr Gln Glu Thr Tyr Ile Asp
            260                 265                 270
Asp Leu Gly His Leu His Val Met Asp Thr Val Phe Leu Ile Gln
        275                 280                 285
Tyr Leu Phe Leu Thr Phe Pro Arg Ile Val Phe Met Leu Gly Phe
    290                 295                 300
Val Val Val Leu Ser Phe Leu Leu Gly Gly Tyr Leu Leu Phe Val
305                 310                 315
Leu Tyr Leu Ala Ala Thr Asn Gln Thr Thr Asn Glu Trp Tyr Arg
                320                 325                 330
Gly Asp Trp Ala Trp Cys Gln Arg Cys Pro Leu Val Ala Trp Pro
            335                 340                 345
Pro Ser Ala Glu Pro Gln Val His Arg Asn Ile His Ser His Gly
        350                 355                 360

Leu Arg Ser Asn Leu Gln Glu Ile Phe Leu Pro Ala Phe Pro Cys
    365                 370                 375
His Glu Arg Lys Lys Gln Glu
380
```

<210> 10
<211> 257
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1659465CD1

<400> 10
Met Ala Ser Lys Ile Gly Ser Arg Arg Trp Met Leu Gln Leu Ile
1 5 10 15
Met Gln Leu Gly Ser Val Leu Leu Thr Arg Cys Pro Phe Trp Gly
20 25 30
Cys Phe Ser Gln Leu Met Leu Tyr Ala Glu Arg Ala Glu Ala Arg
35 40 45
Arg Lys Pro Asp Ile Pro Val Pro Tyr Leu Tyr Phe Asp Met Gly
50 55 60
Ala Ala Val Leu Cys Ala Ser Phe Met Ser Phe Gly Val Lys Arg
65 70 75
Arg Trp Phe Ala Leu Gly Ala Ala Leu Gln Leu Ala Ile Ser Thr
80 85 90
Tyr Ala Ala Tyr Ile Gly Gly Tyr Val His Tyr Gly Asp Trp Leu
95 100 105
Lys Val Arg Met Tyr Ser Arg Thr Val Ala Ile Ile Gly Gly Phe
110 115 120
Leu Val Leu Ala Ser Gly Ala Gly Glu Leu Tyr Arg Arg Lys Pro
125 130 135
Arg Ser Arg Ser Leu Gln Ser Thr Gly Gln Val Phe Leu Gly Ile
140 145 150
Tyr Leu Ile Cys Val Ala Tyr Ser Leu Gln His Ser Lys Glu Asp
155 160 165
Arg Leu Ala Tyr Leu Asn His Leu Pro Gly Gly Glu Leu Met Ile
170 175 180
Gln Leu Phe Phe Val Leu Tyr Gly Ile Leu Ala Leu Ala Phe Leu
185 190 195
Ser Gly Tyr Tyr Val Thr Leu Ala Ala Gln Ile Leu Ala Val Leu
200 205 210
Leu Pro Pro Val Met Leu Leu Ile Asp Gly Asn Val Ala Tyr Trp
215 220 225
His Asn Thr Arg Arg Val Glu Phe Trp Asn Gln Met Lys Leu Leu
230 235 240
Gly Glu Ser Val Gly Ile Phe Gly Thr Ala Val Ile Leu Ala Thr
245 250 255
Asp Gly

<210> 11
<211> 697
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 2120743CD1

<400> 11

```
Met Cys Lys Ser Leu Arg Tyr Cys Phe Ser His Cys Leu Tyr Leu
  1               5                  10                  15
Ala Met Thr Arg Leu Glu Glu Val Asn Arg Glu Val Asn Met His
             20                  25                  30
Ser Ser Val Arg Tyr Leu Gly Tyr Leu Ala Arg Ile Asn Leu Leu
         35                  40                  45
Val Ala Ile Cys Leu Gly Leu Tyr Val Arg Trp Glu Lys Thr Ala
     50                  55                  60
Asn Ser Leu Ile Leu Val Ile Phe Ile Leu Gly Leu Phe Val Leu
 65                  70                  75
Gly Ile Ala Ser Ile Leu Tyr Tyr Tyr Phe Ser Met Glu Ala Ala
             80                  85                  90
Ser Leu Ser Leu Ser Asn Leu Trp Phe Gly Phe Leu Leu Gly Leu
         95                 100                 105
Leu Cys Phe Leu Asp Asn Ser Ser Phe Lys Asn Asp Val Lys Glu
    110                 115                 120
Glu Ser Thr Lys Tyr Leu Leu Leu Thr Ser Ile Val Leu Arg Ile
    125                 130                 135
Leu Cys Ser Leu Val Glu Arg Ile Ser Gly Tyr Val Arg His Arg
    140                 145                 150
Pro Thr Leu Leu Thr Thr Val Glu Phe Leu Glu Leu Val Gly Phe
    155                 160                 165
Ala Ile Ala Ser Thr Thr Met Leu Val Glu Lys Ser Leu Ser Val
    170                 175                 180
Ile Leu Leu Val Val Ala Leu Ala Met Leu Ile Ile Asp Leu Arg
    185                 190                 195
Met Lys Ser Phe Leu Ala Ile Pro Asn Leu Val Ile Phe Ala Val
    200                 205                 210
Leu Leu Phe Phe Ser Ser Leu Glu Thr Pro Lys Asn Pro Ile Ala
    215                 220                 225
Phe Ala Cys Phe Phe Ile Cys Leu Ile Thr Asp Pro Phe Leu Asp
    230                 235                 240
Ile Tyr Phe Ser Gly Leu Ser Val Thr Glu Arg Trp Lys Pro Phe
    245                 250                 255
Leu Tyr Arg Gly Arg Ile Cys Arg Arg Leu Ser Val Val Phe Ala
    260                 265                 270
Gly Met Ile Glu Leu Thr Phe Phe Ile Leu Ser Ala Phe Lys Leu
    275                 280                 285
Arg Asp Thr His Leu Trp Tyr Phe Val Ile Pro Gly Phe Ser Ile
    290                 295                 300
Phe Gly Ile Phe Trp Met Ile Cys His Ile Ile Phe Leu Leu Thr
    305                 310                 315
Leu Trp Gly Phe His Thr Lys Leu Asn Asp Cys His Lys Val Tyr
    320                 325                 330
Phe Thr His Arg Thr Asp Tyr Asn Ser Leu Asp Arg Ile Met Ala
    335                 340                 345
Ser Lys Gly Met Arg His Phe Cys Leu Ile Ser Glu Gln Leu Val
    350                 355                 360
Phe Phe Ser Leu Leu Ala Thr Ala Ile Leu Gly Ala Val Ser Trp
    365                 370                 375
Gln Pro Thr Asn Gly Ile Phe Leu Ser Met Phe Leu Ile Val Leu
    380                 385                 390
```

```
Pro Leu Glu Ser Met Ala His Gly Leu Phe His Glu Leu Gly Asn
                395                 400                 405
Cys Leu Gly Gly Thr Ser Val Gly Tyr Ala Ile Val Ile Pro Thr
                410                 415                 420
Asn Phe Cys Ser Pro Asp Gly Gln Pro Thr Leu Leu Pro Pro Glu
                425                 430                 435
His Val Gln Glu Leu Asn Leu Arg Ser Thr Gly Met Leu Asn Ala
                440                 445                 450
Ile Gln Arg Phe Phe Ala Tyr His Met Ile Glu Thr Tyr Gly Cys
                455                 460                 465
Asp Tyr Ser Thr Ser Gly Leu Ser Phe Asp Thr Leu His Ser Lys
                470                 475                 480
Leu Lys Ala Phe Leu Glu Leu Arg Thr Val Asp Gly Pro Arg His
                485                 490                 495
Asp Thr Tyr Ile Leu Tyr Tyr Ser Gly His Thr His Gly Thr Gly
                500                 505                 510
Glu Trp Ala Leu Ala Gly Gly Asp Thr Leu Arg Leu Asp Thr Leu
                515                 520                 525
Ile Glu Trp Trp Arg Glu Lys Asn Gly Ser Phe Cys Ser Arg Leu
                530                 535                 540
Ile Ile Val Leu Asp Ser Glu Asn Ser Thr Pro Trp Val Lys Glu
                545                 550                 555
Val Arg Lys Ile Asn Asp Gln Tyr Ile Ala Val Gln Gly Ala Glu
                560                 565                 570
Leu Ile Lys Thr Val Asp Ile Glu Glu Ala Asp Pro Pro Gln Leu
                575                 580                 585
Gly Asp Phe Thr Lys Asp Trp Val Glu Tyr Asn Cys Asn Ser Ser
                590                 595                 600
Asn Asn Ile Cys Trp Thr Glu Lys Gly Arg Thr Val Lys Ala Val
                605                 610                 615
Tyr Gly Val Ser Lys Arg Trp Ser Asp Tyr Thr Leu His Leu Pro
                620                 625                 630
Thr Gly Ser Asp Val Ala Lys His Trp Met Leu His Phe Pro Arg
                635                 640                 645
Ile Thr Tyr Pro Leu Val His Leu Ala Asn Trp Leu Cys Gly Leu
                650                 655                 660
Asn Leu Phe Trp Ile Cys Lys Thr Cys Phe Arg Cys Leu Lys Arg
                665                 670                 675
Leu Lys Met Ser Trp Phe Leu Pro Thr Val Leu Asp Thr Gly Gln
                680                 685                 690
Gly Phe Lys Leu Val Lys Ser
                695

<210> 12
<211> 325
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3344986CD1

<400> 12
Met Ser Asp Ser Lys Glu Pro Arg Val Gln Gln Leu Gly Leu Leu
  1               5                  10                  15
```

```
Gly Cys Leu Gly His Gly Ala Leu Val Leu Gln Leu Leu Ser Phe
                20                  25                  30
Met Leu Leu Ala Gly Val Leu Val Ala Ile Leu Val Gln Val Ser
                35                  40                  45
Lys Val Pro Ser Ser Leu Ser Gln Glu Gln Ser Glu Gln Asp Ala
                50                  55                  60
Ile Tyr Gln Asn Leu Thr Gln Leu Lys Ala Ala Val Gly Glu Leu
                65                  70                  75
Ser Glu Lys Ser Lys Leu Gln Glu Ile Tyr Gln Glu Leu Thr Gln
                80                  85                  90
Leu Lys Ala Ala Val Gly Glu Leu Pro Glu Lys Ser Lys Leu Gln
                95                  100                 105
Glu Ile Tyr Gln Glu Leu Thr Arg Leu Lys Ala Ala Val Gly Glu
                110                 115                 120
Leu Pro Glu Lys Ser Lys Leu Gln Glu Ile Tyr Gln Glu Leu Thr
                125                 130                 135
Arg Leu Lys Ala Ala Val Gly Glu Leu Pro Glu Lys Ser Lys Leu
                140                 145                 150
Gln Glu Ile Tyr Gln Glu Leu Thr Arg Leu Lys Ala Ala Val Gly
                155                 160                 165
Glu Leu Pro Asp Gln Ser Lys Gln Gln Gln Ile Tyr Gln Glu Leu
                170                 175                 180
Thr Asp Leu Lys Thr Ala Phe Glu Arg Leu Cys Arg His Cys Pro
                185                 190                 195
Lys Asp Trp Thr Phe Phe Gln Gly Asn Cys Tyr Phe Met Ser Asn
                200                 205                 210
Ser Gln Arg Asn Trp His Asp Ser Val Thr Ala Cys Gln Glu Val
                215                 220                 225
Arg Ala Gln Leu Val Val Ile Lys Thr Ala Glu Glu Gln Asn Phe
                230                 235                 240
Leu Gln Leu Gln Thr Ser Arg Ser Asn Arg Phe Ser Trp Met Gly
                245                 250                 255
Leu Ser Asp Leu Asn Gln Glu Gly Thr Trp Gln Trp Val Asp Gly
                260                 265                 270
Ser Pro Leu Ser Pro Ser Phe Gln Arg Tyr Trp Asn Ser Gly Glu
                275                 280                 285
Pro Asn Asn Ser Gly Asn Glu Asp Cys Ala Glu Phe Ser Gly Ser
                290                 295                 300
Gly Trp Asn Asp Asn Arg Cys Asp Val Asp Asn Tyr Trp Ile Cys
                305                 310                 315
Lys Lys Pro Ala Ala Cys Phe Arg Asp Glu
                320                 325
```

<210> 13
<211> 369
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3576503CD1

<400> 13
Met Lys Ser Pro Phe Tyr Arg Cys Gln Asn Thr Thr Ser Val Glu

```
  1               5                  10                  15
Lys Gly Asn Ser Ala Val Met Gly Gly Val Leu Phe Ser Thr Gly
             20                  25                  30
Leu Leu Gly Asn Leu Leu Ala Leu Gly Leu Leu Ala Arg Ser Gly
             35                  40                  45
Leu Gly Trp Cys Ser Arg Arg Pro Leu Arg Pro Leu Pro Ser Val
             50                  55                  60
Phe Tyr Met Leu Val Cys Gly Leu Thr Val Thr Asp Leu Leu Gly
             65                  70                  75
Lys Cys Leu Leu Ser Pro Val Val Leu Ala Ala Tyr Ala Gln Asn
             80                  85                  90
Arg Ser Leu Arg Val Leu Ala Pro Ala Leu Asp Asn Ser Leu Cys
             95                  100                 105
Gln Ala Phe Ala Phe Phe Met Ser Phe Phe Gly Leu Ser Ser Thr
            110                 115                 120
Leu Gln Leu Leu Ala Met Ala Leu Glu Cys Trp Leu Ser Leu Gly
            125                 130                 135
His Pro Phe Phe Tyr Arg Arg His Ile Thr Leu Arg Leu Gly Ala
            140                 145                 150
Leu Val Ala Pro Val Val Ser Ala Phe Ser Leu Ala Phe Cys Ala
            155                 160                 165
Leu Pro Phe Met Gly Phe Gly Lys Phe Val Gln Tyr Cys Pro Gly
            170                 175                 180
Thr Trp Cys Phe Ile Gln Met Val His Glu Glu Gly Ser Leu Ser
            185                 190                 195
Val Leu Gly Tyr Ser Val Leu Tyr Ser Ser Leu Met Ala Leu Leu
            200                 205                 210
Val Leu Ala Thr Val Leu Cys Asn Leu Gly Ala Met Arg Asn Leu
            215                 220                 225
Tyr Ala Met His Arg Arg Leu Gln Arg His Pro Arg Ser Cys Thr
            230                 235                 240
Arg Asp Cys Ala Glu Pro Arg Ala Asp Gly Arg Glu Ala Ser Pro
            245                 250                 255
Gln Pro Leu Glu Glu Leu Asp His Leu Leu Leu Leu Ala Leu Met
            260                 265                 270
Thr Val Leu Phe Thr Met Cys Ser Leu Pro Val Ile Tyr Arg Ala
            275                 280                 285
Tyr Tyr Gly Ala Phe Lys Asp Val Lys Glu Lys Asn Arg Thr Ser
            290                 295                 300
Glu Glu Pro Glu Arg Pro Pro Ser Leu Ala Ile Ser Ile Cys Asp
            305                 310                 315
Val Asn Cys Gly Pro Leu Asp Ser Tyr His Phe Gln Ile Ser Ser
            320                 325                 330
Ile Ser Asp Ile Phe Ser Gln Asp Phe Ser Leu Asp Leu Leu Gly
            335                 340                 345
Thr Gly Ala Asp Ala Ala Ile Pro Leu Thr Trp Asn Pro Val Cys
            350                 355                 360
Asp Ser Val Phe His Ser Val Val Ser
            365
```

<210> 14
<211> 572
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2355971CB1

<400> 14
gcggagcagc ccgaggcggg gcagcctccc ggagcagcgc cgcgcagagc ccgggacaat 60
ggggccgcgg cggctgctgc tggtggccgc ctgcttcagt ctgtgcggcc cgctgttgtc 120
tgcccgcacc cgggcccgca ggccaggtga gagatgcacg ggaatggggt gcgcgggcgg 180
agggacgccg aggggagact gcgggggtca ctgttgcgac ttctcctcac ccctgcctca 240
gtttcctccg aaagccaaac tggcatttgg gctgagatct ggagtttttt ccagtcacgt 300
ttaggtgggg cgtgccaccc ccttcgttgg cacagccgat gccccttgg actcgatctt 360
ggagggtgca gcccgcctgc aacggggtgt tggatatgga ggaagatgga gcggaagccc 420
ctgggggagc ctgcagtcct gcgttggaat tgtcaacaaa accgtttctt cccaaggacc 480
aacccccaaa aaggaaaagc ttctcaagtt ggtcccaacc aattaaacgt ttcggatctt 540
ttaaaaaaac aaaaaaccaa aaggggcggg cc 572

<210> 15
<211> 517
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2917059CB1

<400> 15
agaaaggaat acccgatgat ggaagtagct cttatggctg gagattgcag gtttatgact 60
gatcctattt gggaagaaca atgatggcag gcattcgagc tttatttatg tacttgtggc 120
tgcagctgga ctgggtgagc agaggagaga gtgtggggct gcatcttcct accctgagtg 180
tccaggaggg tgacaactct attatcaact gtgcttattc aaacagcgcc tcagactact 240
tcatttggta caagcaagaa tctggaaaag gtcctcaatt cattatagac attcgttcaa 300
atatggacaa aaggcaaggc caaagagtca ccgttttatt gaataagaca gtgaaacatc 360
tctctctgca aattgcagct actcaacctg gagactcagc tgtctacttt tgtgcagaga 420
atacacattg ctttccaggc atctgtaacc atcacccaaa cctgagatgg gaggtgaagc 480
agcatccctt tcctttgcaa taaattttag ttataga 517

<210> 16
<211> 2099
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 538181CB1

<400> 16
ggagacctaa acacagtcac catgaagctg ggctgtgtcc tcatggcctg ggccctctac 60
ctttcccttg gtgtgctctg ggtggcccag atgctactgg cagctggatg tcatgccgct 120
gccagttttg agacgctgca gtgtgaggga cctgtctgca ctgaggagag cagctgccac 180
acggaggatg acttgactga tgcaagggaa gctggcttcc aggtcaaggc ctacactttc 240
agtgaaccct tccacctgat tgtgtcctat gactggctga tcctccaagg tccagccaag 300
ccagtttttg aaggggacct gctggttctg cgctgccagg cctggcaaga ctggccactg 360
actcaggtga ccttctaccg agatggctca gctctgggtc cccccgggcc taacagggaa 420

ttctccatca ccgtggtaca aaaggcagac agcgggcact accactgcag tggcatcttc 480
cagagccctg gtcctgggat cccagaaaca gcatctgttg tggctatcac agtccaagaa 540
ctgtttccag cgccaattct cagagctcta ccctcagctg aaccccaagc aggaggcccc 600
atgaccctga gttgtcagac aaagttgccc ctgcagaggt cagctgcccg cctcctcttc 660
tccttctaca aggatggaag gatagtgcaa agcagggggc tctcctcaga attccagatc 720
cccacagctt cagaagatca ctccgggtca tactggtgtg aggcagccac tgaggacaac 780
caagtttgga aacagagccc ccagctagag atcagagtgc agggtgcttc cagctctgct 840
gcacctccca cattgaatcc agctcctcag aaatcagctg ctccaggaac tgctcctgag 900
gaggcccctg ggcctctgcc tccgccgcca accccatctt ctgaggatcc aggcttttct 960
tctcctctgg ggatgccaga tcctcatctg tatcaccaga tgggccttct tctcaaacac 1020
atgcaggatg tgagagtcct cctcggtcac ctgctcatgg agttgaggga attatctggc 1080
caccggaagc ctgggaccac aaaggctact gctgaataga agtaaacagt tcatccatga 1140
tctcacttaa ccaccccaat aaatctgatt ctttattttc tcttcctgtc ctgcacatat 1200
gcataagtac ttttacaagt tgtcccagtg ttttgttaga ataatgtagt taggtgagtg 1260
taaataaatt tatataaagt gagaattaga gtttagctat aattgtgtat tctctcttaa 1320
cacaacagaa ttctgctgtc tagatcagga atttctatct gttatatcga ccagaatgtt 1380
gtgatttaaa gagaactaat ggaagtggat tgaatacagc agtctcaact gggggcaatt 1440
ttgccccca gaggacattg ggaaatgttt ggagacattt tggtcattat acttgggggg 1500
ttgggggatg gtgggatgtg tgtgctactg gcatccagta aatagaagcc aggggtgccg 1560
ctaaacatcc tataatgcac agggcagtac cccacaacga aaaataatct ggcccaaaat 1620
gtcagttgta ctgagtttga gaaaccccag cctaatgaaa ccctaggtgt tgggctctgg 1680
aatgggactt tgtcccttct aattattatc tctttccagc ctcattcagc tattcttact 1740
gacataccag tctttagctg gtgctatggt ctgttcttta gttctagttt gtatcccctc 1800
aaaagccatt atgttgaaat cctaatcccc aaggtgatgg cattaagaag tgggcctttg 1860
ggaagtgatt agatcaggag tgcagagccc tcatgattag gattagtgcc cttatttaaa 1920
aaggccccag agagctaact caccccttcca ccatatgagg acgtggcaag aagatgacat 1980
gtatgagaac caaaaaacag tgtcgccaaa caccgactct gtcgttgcct tgatcttgaa 2040
cttccagcct ccagaactat gagaaataaa attctgttgt ttgtaagcta aaaaaaaaa 2099

<210> 17
<211> 770
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1368760CB1

<400> 17
agtgtggggt ggaaaggaag agtgagcgag agcaagttaa ggggaggggg tgtaagagcc 60
agcgaattct ttttcttttt ctattattat tttgacgact cctgagttgc gcccatgctc 120
ttgtcagctt cgttttaggc gtacatggcc aggcagaaga aaatggggca aagcgtgctc 180
cgggcggtct tctttttagt cctggggctt ttgggtcatt ctcacggagg attccccaac 240
accatcagca taggtggact tttcatgaga aacacagtgc aggagcacag cgctttccgc 300
tttgccgtgc agttatacaa caccaaccag aacaccaccg agaagccctt ccatttgaat 360
taccacgtag atctcttgga ttcctccaat agtttttccg tgacaaatgc tttctgctcc 420
cagttctcga gaggggtgta tgccatcttt ggattctatg accagatgtc aatgaacacc 480
ctgacctcct tctgtggggc cctgcacaca tcctttgtta cgcctagctt ccccactgac 540
gcagatgtgc agtttgtcat ccagatgcgc ccagccttga agggcgctat tctgagtctt 600
ctgggtcatt acaagtggga gaagtttgtg tacctctatg acacagaacg aggtaagaag 660
aggcacctgc tctgctcttt agatattcat gtaattgtgt tcaaacttcc tcagcttatg 720
tgccctttgc ttccaataaa taaaatctaa ttctgtttta aaaaaaaaaa 770

<210> 18
<211> 572
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1670669CB1

<400> 18
gcctggtcgg gaagggcctg gtcagctgcg tccggcggag gcagctgctg acccagctgt 60
ggactgtgcc ggtggcgggg gacggagggg caggagccct gggctccccg tgggcgggag 120
ctgtatcatg gaccacctcg gggcgtccct ctggccccag gtcggctccc tttgtctcct 180
gctcgctggg gccgcctggg cgcccccgcc taacctcccg gaccccaagt tcgagagcaa 240
agcggccttg ctggcggccc gggggcccga agagcttctg tgcttcaccg agcgggttgg 300
aggacttggg atgagccatg ggaagctgtg tcgcctgcac caggctccca cggcgcgtgg 360
tggcggtgcg cttctggtgt gcgctgccta caggcgacac gtcgagcttc gtgcccctag 420
agttgggcgt cacagcagcc tccggcgctc cgcgatatca ccgtgtcatc cacatcaatg 480
aagtagtgct ccgagacggc cccgtggggc tggtggcgcg ggttggctga cgagagcggg 540
cagctaggtg tgcgctggct cccggggggt gg 572

<210> 19
<211> 1795
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2851578CB1

<400> 19
ctcagtaccc cgaggcggat ggaaggaggg agggctgcag ggttcccctt ggcctgcaaa 60
caggaacaca gggtgtttct cagtggctgc gagaatgctg atgaaaaccc caggatgttg 120
tgtcaccgtg gtggccagct gatagtgcca atcatcccac tttgccctga gcactcctgc 180
aggggtagaa gactccagaa ccttctctca ggcccatggc ccaagcagcc catggaactt 240
cataacctga gctctccatc tccctctctc tcctcctctg ttctccctcc ctccttctct 300
ccctcaccct cctctgctcc ctctgccttt accactgtgg gggggtcctc tggagggccc 360
tgccacccca cctcttcctc gctggtgtct gccttcctgg caccaatcct ggccctggag 420
tttgtcctgg gcctggtggg gaacagtttg gccctcttca tcttctgcat ccacacgcgg 480
ccctggacct ccaacacggt gttcctggtc agcctggtgg ccgctgactt cctcctgatc 540
agcaacctgc ccctccgcgt ggactactac ctcctccatg agacctggcg ctttggggct 600
gctgcctgca aagtcaacct cttcatgctg tccaccaacc gcacggccag cgttgtcttc 660
ctcacagcca tcgcactcaa ccgctacctg aaggtggtgc agccccacca cgtgctgagc 720
cgtgcttccg tgggggcagc tgcccgggtg gccgggggac tctgggtggg catcctgctc 780
ctcaacgggc acctgctcct gagcaccttc tccggcccct cctgcctcag ctacagggtg 840
ggcacgaagc cctcggcctc gctccgctgg caccaggcac tgtacctgct ggagttcttc 900
ctgccactgg cgctcatcct ctttgctatt gtgagcattg ggctcaccat ccggaaccgt 960
ggtctgggcg ggcaggcagg cccgcagagg gccatgcgtg tgctggccat ggtggtggcc 1020
gtctacacca tctgcttctt gcccagcatc atctttggca tggcttccat ggtggctttc 1080
tggctgtccg cctgccgctc cctggacctc tgcacacagc tcttccatgg ctccctggcc 1140
ttcacctacc tcaacagtgt cctggacccc gtgctctact gcttctctag ccccaacttc 1200
ctccaccaga gccgggcctt gctgggcctc acgcggggcc ggcagggccc agtgagcgac 1260
gagagctcct accaacctc caggcagtgg cgctaccggg aggcctctag gaaggcggag 1320
gccataggga agctgaaagt gcagggcgag gtctctctgg aaaaggaagg ctcctcccag 1380

ggctgagggc cagctgcagg gctgcagcgc tgtgggggta agggctgccg cgctctggcc 1440
tggagggaca aggccagcac acggtgcctc aaccaactgg acaagggatg gcggcagacc 1500
aggggccagg ccaaagcact ggcaggactc aggtgggtgg cagggagaga aacccaccta 1560
ggccttctcag tgtgtccagg atggcattcc cagaatgcag gggagagcag gatgccgggt 1620
ggaggagaca ggcaaggtgc cgttggcaca ccagctcaga caggggcctg cgcagctgca 1680
ggggacagac gccaatcact gtcacagcag agtcaccttа gaaattggac agctgcatgt 1740
tctgtgctct ccagtttgtc ccttccaata ttaataaact tccctttaa atata 1795

<210> 20
<211> 2053
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3393757CB1

<400> 20
cccaccctca tcggcctccc aaagtgctgg aattagaggc gtgatcaccg tgcccagccg 60
ccaatgccat cttcatcccc cagatagaca gtctctagga tctgttccct ggggctgagc 120
ggttggagtc ttcatgcggg ccctctggcc catggctcac taggtctgtg tccacatccc 180
tccagagcat cttacaggcg ctggatgagc tgctggaggc ccctggggac ctggagaccc 240
tgccccgctt acagcagcac tgtgtggcca gtcacctgct ggatggccta gaggatgtcc 300
tcagaggcct gagcaagaac ctttccaatg ggctgttgaa cttcagttat cctgcaggca 360
cagaattgtc cctggaggtg cagaagcaag tagacaggag tgtcaccttg agacagaatc 420
aggcagtgat gcagctcgac tggaatcagg cacagaaatc tggtgaccca ggccccttctg 480
tggtgggcct tgtctccatt ccagggatgg gcaagttgct ggctgaggcc cctctggtcc 540
tggaacctga gaagcagatg cttctgcatg agacacacca gggcttgctg caggacggct 600
cccccatcct gctctcagat gtgatctctg ccttttctgag caacaacgac acccaaaacc 660
tcagctcccc agttaccttc accttctccc accgttcagt gatcccgaga cagaaggtgc 720
tctgtgtctt ctgggagcat ggccagaatg gatgtggtca ctgggccacc acaggctgca 780
gcacaatagg caccagagac accagcacca tctgccgttg cacccacctg agcagctttg 840
ccgtcctcat ggcccactac gatgtgcagg aggaggatcc cgtgctgact gtcatcacct 900
acatggggct gagcgtctct ctgctgtgcc tcctcctggc ggccctcact tttctcctgt 960
gtaaagccat ccagaacacc agcacctcac tgcatctgca gctctcgctc tgcctctttcc 1020
tggcccacct cctcttcctc gtggcaattg atcaaaccgg acacaaggtg ctgtgctcca 1080
tcatcgccgg taccttgcac tatctctacc tggccacctt gacctggatg ctgctggagg 1140
ccctgtacct cttcctcact gcacggaacc tgacggtggt caactactca agcatcaaca 1200
gattcatgaa gaagctcatg ttccctgtgg gctacggagt cccagctgtg acagtggcca 1260
tttctgcagc ctccaggcct caccttttatg gaacaccttc ccgctgctgg ctccaaccag 1320
aaaagggatt tatatggggc ttccttggac ctgtctgcgc catcttctct gtgaatttag 1380
ttctctttct ggtgactctc tggattttga aaaacagact ctcctccctc aatagtgaag 1440
tgtccaccct ccggaacaca aggatgctgg catttaaagc gacagctcag ctgttcatcc 1500
tgggctgcac gtggtgtctg ggcatcttgc aggtgggtcc ggctgcccgg gtcatggcct 1560
acctcttcac catcatcaac agcctgcagg gtgtcttcat cttcctggtg tactgcctcc 1620
tcagccagca ggtccgggag caatatggga aatggtccaa agggatcagg aaattgaaaa 1680
ctgagtctga gatgcacaca ctctccagca gtgctaaggc tgacacctcc aaacccagca 1740
cggtaagatc acgcattgct ccagagcact tcactaaccg acccacctga ggagcatgtg 1800
cctatcacac aaggaaacct gggaatacag caggcaatgc cctagaaagg ctcgcatctg 1860
agtacgcctt gactcattaa ccattagcaa tgatctcagt ttaaatgttt ttttttaatc 1920
agtcatagcc tgtcatcccg gcatcactgt catcccagca tttgggaggc ctaggcaaga 1980
gggtcacctg aggccaggag tgcaagatga ccctgggcaa catagcaaga tcccatctct 2040
acaaaaaaaa aaa 2053

<210> 21
<211> 1500
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 312256CB1

<400> 21
caggtgcccg acatggcgag tgtagtgctg ccgagcggat cccagtgtgc ggcggcagcg 60
gcggcggcgg cgcctcccgg gctccggctc cggcttctgc tgttgctctt ctccgccgcg 120
gcactgatcc ccacaggtga tgggcagaat ctgtttacga aagacgtgac agtgatcgag 180
ggagaggttg cgaccatcag ttgccaagtc aataagagtg acgactctgt gattcagcta 240
ctgaatccca acaggcagac catttatttc agggacttca ggcctttgaa ggacagcagg 300
tttcagttgc tgaatttttc tagcagtgaa ctcaaagtat cattgacaaa cgtctcaatt 360
tctgatgaag gaagatactt ttgccagctc tataccgatc ccccacagga aagttacacc 420
accatcacag tcctggtccc accacgtaat ctgatgatcg atatccagaa agacactgcg 480
gtggaaggtg aggagattga agtcaactgc actgctatgg ccagcaagcc agccacgact 540
atcaggtggt tcaaagggaa cacagagcta aaaggcaaat cggaggtgga agagtggtca 600
gacatgtaca ctgtgaccag tcagctgatg ctgaaggtgc acaaggagga cgatggggtc 660
ccagtgatct gccaggtgga gcaccctgcg gtcactggaa acctgcagac ccagcggtat 720
ctagaagtac agtataagcc tcaagtgcac attcagatga cttatcctct acaaggctta 780
acccgggaag gggacgcgct tgagttaaca tgtgaagcca tcgggaagcc ccagcctgtg 840
atggtaactt gggtgagagt cgatgatgaa atgcctcaac acgccgtact gtctgggccc 900
aacctgttca tcaataacct aaacaaaaca gataatggta cataccgctg tgaagcttca 960
aacatagtgg ggaaagctca ctcggattat atgctgtatg tatacgatcc ccccacaact 1020
atccctcctc ccacaacaac caccaccacc accaccacca ccaccaccac catccttacc 1080
atcatcacag attcccgagc aggtgaagaa ggctcgatca gggcagtgga tcatgccgtg 1140
atcggtggcg tcgtggcggt ggtggtgttc gccatgctgt gcttgctcat cattctgggg 1200
cgctattttg ccagacataa aggtacatac ttcactcatg aagccaaagg agccgatgac 1260
gcagcagacg cagacacagc tataatcaat gcagaaggag gacagaacaa ctccgaagaa 1320
aagaaagagt acttcatcta gatcagcctt tttgtttcaa tgaggtgtcc aactggccct 1380
atttagatga taaagagaca gtgatattgg aacttgcgag aaattcgtgt gttttttat 1440
gaatgggtgg aaaggtgtga gactgggaag gcttgggatt tgctgtgtaa aaaaaaaaaa 1500

<210> 22
<211> 1449
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1615704CB1

<400> 22
gtcacgagcc cgcaggaagt ctcgtatcgc gcccgggagg cgccggagcc cagcggctgg 60
cgccagatcc aggctcctgg aagaaccatg tccggcagct actggtcatg ccaggcacac 120
actgctgccc aagaggagct gctgtttgaa ttatctgtga atgttgggaa gaggaatgcc 180
agagctgccg gctgaaaatt acccaaccaa gagaaatctg caggatggac tttctggtcc 240
tcttcttgtt ctacctggct tcggtgctga tgggtcttgt tcttatctgc gtctgctcga 300
aaacccatag cttgaaaggc ctggccaggg gaggagcaca gatattttcc tgtataattc 360
cagaatgtct tcagagagcc gtgcatggat tgcttcatta ccttttccat acgagaaacc 420
acaccttcat tgtcctgcac ctggtcttgc aagggatggt ttatactgag tacacctggg 480

aagtatttgg ctactgtcag gagctggagt tgtccttgca ttacctttctt ctgccctatc 540
tgctgctagg tgtaaaccctg tttttttca ccctgacttg tggaaccaat cctggcatta 600
taacaaaagc aaatgaatta ttatttcttc atgtttatga atttgatgaa gtgatgtttc 660
caaagaacgt gaggtgctct acttgtgatt taaggaaacc agctcgatcc aagcactgca 720
gtgagtgtgg ctctcgtgac tccagcggca cctccaacag cacatgtgtg ggcttcgtct 780
gtgagggaat gtttcctgaa tccgaaagca gagccagttc acccccagat atggtgtgtg 840
taacctggtg tgtgcaccgt ttcgaccatc actgtgtttg ggtgaacaac tgcatcgggg 900
cctggaacat caggtacttc ctcatctacg tcttgacctt gacggcctcg gctgccaccg 960
tcgccattgt gagcaccact tttctggtcc acttggtggt gatgtcagat ttataccagg 1020
agacttacat cgatgacctt ggacacctcc atgttatgga cacggtcttt cttattcagt 1080
acctgttcct gacttttcca cggattgtct tcatgctggg ctttgtcgtg gttctgagct 1140
tcctcctggg tggctacctg ttgtttgtcc tgtatctggc ggccaccaac cagactacta 1200
acgagtggta cagaggtgac tgggcctggt gccagcgttg tccccttgtg gcctggcctc 1260
cgtcagcaga gccccaagtc caccggaaca ttcactccca tgggcttcgg agcaaccttc 1320
aagagatctt tctacctgcc tttccatgtc atgagaggaa gaaacaagaa tgacaagtgt 1380
atgactgcct ttgagctgta gttcccgttt atttacacat gtggatcctc gttttccaaa 1440
aaaaaaaaa 1449

<210> 23
<211> 1587
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1659465CB1

<400> 23
tgtcccatcc ccgcgcctgc ttctctgact ggggtgaggc cgcagcggac tgccctttcc 60
caagatggcg tcgaagatag gttcgagacg gtggatgttg cagctgatca tgcagttggg 120
ttcggtgctg ctcacacgct gccccttttg gggctgcttc agccagctca tgctgtacgc 180
tgagagggct gaggcacgcc ggaagcccga catcccagtg ccttacctgt atttcgacat 240
gggggcagcc gtgctgtgcg ctagtttcat gtcctttggc gtgaagcggc gctggttcgc 300
gctgggggcc gcactccaat tggccattag cacctacgcc gcctacatcg ggggctacgt 360
ccactacggg gactggctga aggtccgtat gtactcgcgc acagttgcca tcatcggcgg 420
ctttcttgtg ttggccagcg gtgctgggga gctgtaccgc cggaaacctc gcagccgctc 480
cctgcagtcc accggccagg tgttcctggg tatctacctc atctgtgtgg cctactcact 540
gcagcacagc aaggaggacc ggctggcgta tctgaaccat ctcccaggag gggagctgat 600
gatccagctg ttcttcgtgc tgtatggcat cctggccctg gcctttctgt caggctacta 660
cgtgaccctc gctgcccaga tcctggctgt actgctgccc cctgtcatgc tgctcattga 720
tggcaatgtt gcttactggc acaacacgcg gcgtgttgag ttctggaacc agatgaagct 780
ccttggagag agtgtgggca tcttcggaac tgctgtcatc ctggccactg atggctgagt 840
tttatggcaa gaggctgaga tgggcacagg gagccactga gggtcaccct gccttcctcc 900
ttgctggccc agctgctgtt tatttatgct ttttggtctg tttgtttgat cttttgcttt 960
tttaaaattg ttttttgcag ttaagaggca gctcatttgt ccaaatttct gggctcagcg 1020
cttgggaggg caggagccct ggcactaatg ctgtacaggt ttttttcctg ttaggagagc 1080
tgaggccagc tgcccactga gtctcctgtc cctgagaagg gagtatggca gggctgggat 1140
gcggctactg agagtgggag agtgggagac agaggaagga agatggagat tggaagtgag 1200
caaatgtgaa aaattcctct ttgaacctgg cagatgcagc taggctctgc agtgctgttt 1260
ggagactgtg agagggagtg tgtgtgttga cacatgtgga tcaggcccag gaagggcaca 1320
ggggctgagc actacagaag tcacatgggt tctcagggta tgccaggggc agaaacagta 1380
ccggctctct gtcactcacc ttgagagtag agcagaccct gttctgctct gggctgtgaa 1440
ggggtggagc aggcagtggc cagctttgcc cttcctgctg tctctgtttc tagctccatg 1500
gttggcctgg tgggggtgga gttcccctccc aaacaccaga ccacacagtc ctccaaaaat 1560

aaacatttta tatagaaaaa aaaaaaa

<210> 24
<211> 3928
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2120743CB1

<400> 24
ccctcgagcc gtaccgtcgc ggatttcggc ggcggaaaca tggcggtcgc ggccgggccg 60
gtaacggaga aagtttacgc cgacactggc ctgtattagc gcgtatggcc tcgggccctc 120
gttccccaag gcgtgccgcc tccctgttct cagtcgcagg ctgaagcctt gtctgctctc 180
ctcctttttg gtttggtttt ggaactgact ccgagggttg ggagagcgcg ttggtggcga 240
cggccgagtc agatcactat aaacaaaatt tccacaagag aaaatgttga aataggagtt 300
gcggatacat tggatatact ggatgaaata caagcggtta attttgtaac gtgagggaaa 360
agcccacatt gctggttaca tgtgtaaatc actgcgttat tgctttagtc attgtctcta 420
tttagcaatg acaagactgg aagaagtaaa tagagaagtg aacatgcatt cttcagtgcg 480
gtatcttggc tatttagcca gaatcaattt attggttgct atatgcttag gtctatacgt 540
aagatgggaa aaaacagcaa attccttaat tttggtaatt tttattcttg gtctttttgt 600
tcttggaatc gccagcatac tctattacta tttttcaatg gaagcagcaa gtttaagtct 660
ctccaatctt tggtttggat tcttgcttgg cctcctatgt tttcttgata attcatcctt 720
taaaaatgat gtaaaagaag aatcaaccaa atatttgctt ctaacatcca tagtgttaag 780
gatattgtgc tctctggtgg agagaatttc tggttatgtc cgtcatcggc ccactttact 840
aaccacagtt gaatttctgg agcttgttgg atttgccatt gccagcacaa ctatgttggt 900
ggagaagtct ctgagtgtca ttttgcttgt tgtagctctg gctatgctga ttattgatct 960
gagaatgaaa tctttcttag ctattccaaa cttagttatt tttgcagttt tgttattttt 1020
ttcctcattg gaaactccca aaaatccgat tgcttttgcg tgttttttta tttgcctgat 1080
aactgatcct ttccttgaca tttattttag tggactttca gtaactgaaa gatggaaacc 1140
ctttttgtac cgtggaagaa tttgcagaag actttcagtc gtttttgctg gaatgattga 1200
gcttacattt tttattcttt ccgcattcaa acttagagac actcacctct ggtattttgt 1260
aatacctggc ttttccattt ttggaatttt ctggatgatt tgtcatatta tttttctttt 1320
aactctttgg ggattccata ccaaattaaa tgactgccat aaagtatatt ttactcacag 1380
gacagattac aatagccttg atagaatcat ggcatccaaa gggatgcgcc atttttgctt 1440
gatttcagag cagttggtgt tctttagtct tcttgcaaca gcgattttgg gagcagtttc 1500
ctggcagcca acaaatggaa ttttcttgag catgtttcta atcgttttgc cattggaatc 1560
catggctcat gggctcttcc atgaattggg taactgttta ggaggaacat ctgttggata 1620
tgctattgtg attcccacca acttctgcag tcctgatggt cagccaacac tgcttccccc 1680
agaacatgta caggagttaa atttgaggtc tactggcatg ctcaatgcta tccaaagatt 1740
ttttgcatat catatgattg agacctatgg atgtgactat tccacaagtg gactgtcatt 1800
tgatactctg cattccaaac taaaagcttt cctcgaactt cggacagtgg atggacccag 1860
acatgatacg tatattttgt attacagtgg gcacacccat ggtacaggag agtgggctct 1920
agcaggtgga gatacactac gccttgacac acttatagaa tggtggagag aaaagaatgg 1980
ttccttttgt tcccggctta ttatcgtatt agacagcgaa aattcaaccc cttgggtgaa 2040
agaagtgagg aaaattaatg accagtatat tgcagtgcaa ggagcagagt tgataaaaac 2100
agtagatatt gaagaagctg acccgccaca gctaggtgac tttacaaaag actgggtaga 2160
atataactgc aactccagta ataacatctg ctggactgaa aagggacgca cagtgaaagc 2220
agtatatggt gtgtcaaaac ggtggagtga ctacactctg catttgccaa cgggaagcga 2280
tgtggccaag cactggatgt tacactttcc tcgtattaca tatcccctag tgcatttggc 2340
aaattggtta tgcggtctga acctttttttg gatctgcaaa acttgtttta ggtgcttgaa 2400
aagattaaaa atgagttggt ttcttcctac tgtgctggac acaggacaag gcttcaaact 2460
tgtcaaatct taatttggac cccaaagcgg gatattaata agcactcata ctaccaatta 2520

tcactaactt gccatttttt gtatgctgta tttttatttg tggaaaatac cttgctactt 2580
ctgtagctgc tctcactttg tcttttctta agtaattatg gtatatataa ggcgttggga 2640
aaaaacattt tataatgaaa gtatgtaggg agtcaaatgc ttactgtaaa tgcataagag 2700
acgttaaaaa taacactgca ctttcaggaa tgtttgctta tggtcctgat tagaaagaaa 2760
cagttgtcta tgctctgcaa tggtcaatga tgaattacta atgccttatt ttctaggcat 2820
ataataatag tttagagaat gtagaccaga taaatttgtt tactgtttta agaaaactac 2880
cagtttactt acagaagatt ctttttttca aacagtaggt ttcatccaag accatttgaa 2940
gaactgcaaa ctctttctct tagaaaagaa agagggcagc ctaaaataaa cgcaaaattt 3000
gcttatactc catcacattc agatgtcttg gttgtgactt attaccagtg tggcagagaa 3060
cccaagttac attttagatc aaaatattct ttatgtaggt attgttaaaa ggctagagcc 3120
tacaagttgc tcttccatgc gttggtcagg gggccctgaa aacactggta atattaagag 3180
tctttcttcag ggtaacttaa tgttttctta atgaacagtg tttccagcta caaattcttc 3240
caataaattg tcttcctttt tgaaaagtac tctcatagaa gaaatttagc aatttctcgt 3300
tgactgactc agtctatttt aagtattcag aaaagatttt gatccccatt gagttaatgc 3360
tctgccttga aaattatttt tctgatcctt gttagtgata acatttttct tctactgaag 3420
gtcagaggat aggaaacaag tatttctctt ctggtataca tgtaatgtat tctgtaaaaa 3480
agtattcata ttggcaattt tagttaggca taatattgtg gttgtaattt ttaaaactta 3540
gtgttttgtc tgattaaagc aggcactgat cagggtatct cctaagaggt aattcacttc 3600
ttattccttt ccaataatta ttacattcta aattttcatc tatgagaaat aacaaacaag 3660
aagggaatag aattaaattg gggtataatc taatcttcat tgtttaaatg gtttgccttc 3720
tcaccattga agccattttt ttatagcctc agaaagagga aataatgcct ccaccatttt 3780
ctacctggtg acttgaaaat tgaactttta agttaggaag aagttagagt cagggaactt 3840
gtataccact atctatgcag cattgttata gtctgattat ttctgtgttt tgaatatgat 3900
tttcctaatg ctctaaataa aatttttc 3928

<210> 25
<211> 1542
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3344986CB1

<400> 25
atgtaccacc gcgtccgcat gtcatgagtg gaggagtcct tctccctgtc aaccccagac 60
catcccccaa cacctccctg aaattcctgc aaggtcaggc cgatctcagg ctctgactcc 120
tttcactagc ctttctggtg atgtgatctt acccagcttc ctgtttgtct tcctgagaga 180
cagtagattt agaaagtgag gatcagaggg tggaaaataa aagctgtggt ccccaggagt 240
cctgaacatc tgggggcagc gggaaaacat gagtgactcc aaggaaccaa gggtgcagca 300
gctgggcctc ctggggtgtc ttggccatgg cgccctggtg ctgcaactcc tctccttcat 360
gctcttggct ggggtcctgg tggccatcct tgtccaagtg tccaaggtcc ccagctccct 420
aagtcaggaa caatccgagc aagacgcaat ctaccagaac ctgacccagc ttaaagctgc 480
agtgggtgag ctctcagaga aatccaagct gcaggagatc taccaggagc tgacccagct 540
gaaggctgca gtgggtgagt tgccagagaa atccaagctg caggagatct accaggagct 600
gacccggctg aaggctgcag tgggtgagtt gccagagaaa tccaagctgc aggagatcta 660
ccaggagctg acccggctga aggctgcagt gggtgagttg ccagagaaat ccaagctgca 720
ggagatctac caggagctga cccggctgaa ggctgcagtg ggtgagttgc cagaccagtc 780
caagcagcag caaatctacc aggagctgac cgatttgaag actgcatttg aacgcctgtg 840
ccgccactgt cccaaggact ggacattctt ccaaggaaac tgttacttca tgtctaactc 900
ccagcggaac tggcacgact ccgtcaccgc ctgccaggaa gtgagggccc agctcgtcgt 960
aatcaaaact gctgaggagc agaacttcct acagctgcag acttccagga gtaaccgctt 1020
ctcctggatg ggactttcag acctaaatca ggaaggcacg tggcaatggg tggacggctc 1080
acctctgtca cccagcttcc agcggtactg gaacagtgga gaacccaaca atagcgggaa 1140

tgaagactgt gcggaattta gtggcagtgg ctggaacgac aatcgatgtg acgttgacaa 1200
ttactggatc tgcaaaaagc ccgcagcctg cttcagagac gaatagttgt ttccctgcta 1260
gcctcagcct ccattgtggt atagcagaac ttcacccact tctacacccc gtgcacccct 1320
ttgactgggg acttgctggt tgaaggagct catcttgcag gctggaagca ccagggaatt 1380
aattccccca gtcaaccaat ggcatccaga gagggcatgg aggctccata caacctcttc 1440
caccccaca tctttctttg tcctatacat gtcttccatt tggctgtttc tgagttgtag 1500
cctttataat aaagtggtaa atgttgtaac tgcaaaaaaa aa 1542

<210> 26
<211> 1264
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3576503CB1

<400> 26
cggacttttt ctgtggcgca gcttctccgc ccgagccgcg cgcggagctg ccgggggctc 60
cttagcaccc gggcgccggg gccctcgccc ttccgcagcc ttcactccag ccctctgctc 120
ccgcacgcca tgaagtcgcc gttctaccgc tgccagaaca ccacctctgt ggaaaaaggc 180
aactcggcgg tgatgggcgg ggtgctcttc agcaccggcc tcctgggcaa cctgctggcc 240
ctggggctgc tggcgcgctc ggggctgggg tggtgctcgc ggcgtccact gcgcccgctg 300
ccctcggtct tctacatgct ggtgtgtggc ctgacggtca ccgacttgct gggcaagtgc 360
ctcctaagcc cggtggtgct ggctgcctac gctcagaacc ggagtctgcg ggtgcttgcg 420
cccgcattgg acaactcgtt gtgccaagcc ttcgccttct tcatgtcctt ctttgggctc 480
tcctcgacac tgcaactcct ggccatggca ctggagtgct ggctctccct agggcaccct 540
ttcttctacc gacggcacat caccctgcgc ctgggcgcac tggtggcccc ggtggtgagc 600
gccttctccc tggctttctg cgcgctacct ttcatgggct tcgggaagtt cgtgcagtac 660
tgccccggca cctggtgctt tatccagatg gtccacgagg agggctcgct gtcggtgctg 720
gggtactctg tgctctactc cagcctcatg gcgctgctgg tcctcgccac cgtgctgtgc 780
aacctcggcg ccatgcgcaa cctctatgcg atgcaccggc ggctgcagcg gcacccgcgc 840
tcctgcacca gggactgtgc cgagccgcgc gcggacggga gggaagcgtc ccctcagccc 900
ctggaggagc tggatcacct cctgctgctg gcgctgatga ccgtgctctt cactatgtgt 960
tctctgcccg taatttatcg cgcttactat ggagcattta aggatgtcaa ggagaaaaac 1020
aggacctctg aagaaccaga gagacctccg agccttgcga tttctatctg tgatgtcaat 1080
tgtggaccct tggattctta tcatttttcag atctccagta tttcggatat tttttcacaa 1140
gatttttcat tagacctctt aggtacagga gccgatgcag caattccact aacatggaat 1200
ccagtctgtg acagtgtttt tcactctgtg gtaagctgag gagtgtctga catgcggtgg 1260
tggt 1264

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, C12Q 1/68, A61K 38/17, C07K 16/18 — A3

(11) International Publication Number: **WO 00/28032**

(43) International Publication Date: 18 May 2000 (18.05.00)

(21) International Application Number: PCT/US99/26742

(22) International Filing Date: 12 November 1999 (12.11.99)

(30) Priority Data:

09/191,280	12 November 1998 (12.11.98)	US
Not furnished	12 November 1998 (12.11.98)	US
09/206,647	7 December 1998 (07.12.98)	US
Not furnished	7 December 1998 (07.12.98)	US
60/123,404	8 March 1999 (08.03.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	09/191,280 (CIP)
Filed on	12 November 1998 (12.11.98)
US	Not furnished (CIP)
Filed on	12 November 1998 (12.11.98)
US	09/206,647 (CIP)
Filed on	7 December 1998 (07.12.98)
US	Not furnished (CIP)
Filed on	7 December 1998 (07.12.98)
US	60/123,404 (CIP)
Filed on	8 March 1999 (08.03.99)

(71) Applicant (*for all designated States except US*): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Moutain View, CA 94043 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). AU-YOUNG, Janice [US/US]; 1419 Kains Avenue, Berkeley, CA 94702 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
16 November 2000 (16.11.00)

(54) Title: HUMAN CELL SURFACE RECEPTOR PROTEINS

(57) Abstract

The invention provides human cell surface receptor proteins (HCSRP) and polynucleotides which identify and encode HCSRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of HCSRP.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

Internat al Application No
PCT/US 99/26742

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/705 C12Q1/68 A61K38/17 C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 18456 A (UNIVERSITY OF CALIFORNIA) 7 May 1998 (1998-05-07) the whole document, especially residues 1-29 in SEQ ID NO:7	1-16,19
X	WO 92 14750 A (UNIVERSITY OF CALIFORNIA ;COR THERAPEUTICS INC) 3 September 1992 (1992-09-03) the whole document, especially nucleotides 166-312 in Figure 1A	1-16,19
X	WO 94 05695 A (UNIVERSITY OF NEW YORK) 17 March 1994 (1994-03-17) the whole document,especially sequence 31 in figure 8B(2)	1,2,15

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 March 2000	25. 07. 00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	CUPIDO, M

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	EMBL/GenBank/DDBJ databases SEQ ID AC008515:Homo sapiens chromosome 5 clone CTC-455F7 4 August 1999 DOE Joint Genome Institute XP002132666 compare nucleotides 40760-41260 with nucleotides 516-1 in SEQ ID NO:14 -----	3-6,9-11

INTERNATIONAL SEARCH REPORT

In ational application No.
PCT/US 99/26742

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claim 19 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the composition.

2. [X] Claims Nos.: 17,18 and 20
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See additional sheets

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-20 (all partly)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-20 (all partly)

Polypeptide comprising an amino acid sequence of SEQ ID NO:1, fragments and variants thereof, polynucleotides encoding them, methods of detecting related sequences, vectors and cells comprising said sequences and their use to produce a polypeptide, corresponding pharmaceutical compositions and antibodies.

2. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:2

3. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:3

4. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:4

5. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:5

6. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:6

7. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:7

8. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:8

9. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:9

10. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:10

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

11. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:11

12. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:12

13. Claims: 1-20 (all partly)

Idem as subject 1 but limited to SEQ ID NO:13

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17,18 and 20

Claims 17, 18 and in part 20 refer to an antagonist and agonist of the polypeptides without giving a true technical characterisation. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Articles 5 and 6 PCT). No meaningful search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

...rmation on patent family members

Interr ...nal Application No
PCT/US 99/26742

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9818456 A	07-05-1998	US 5892014 A	06-04-1999
		EP 0948323 A	13-10-1999
WO 9214750 A	03-09-1992	US 5256766 A	26-10-1993
		US 5688768 A	18-11-1997
		AU 665752 B	18-01-1996
		AU 1456892 A	15-09-1992
		CA 2104394 A	20-08-1992
		EP 0572553 A	08-12-1993
		JP 6508742 T	06-10-1994
		NZ 241666 A	26-07-1995
		US 6024936 A	15-02-2000
		US 5759994 A	02-06-1998
		US 5856448 A	05-01-1999
		US 5849507 A	15-12-1998
		US 5798248 A	25-08-1998
WO 9405695 A	17-03-1994	AU 4855393 A	29-03-1994
		US 5508384 A	16-04-1996